

BUNDESREPUBLIK DEUTSCHLAND



Prioritätsbescheinigung über die Einreichung einer Patentanmeldung

Aktenzeichen: 199 32 924.9

Anmeldestag: 14. Juli 1999

Anmelder/Inhaber: BASF Aktiengesellschaft, Ludwigshafen/DE

Bezeichnung: Corynebacterium glutamicum genes encoding
proteins involved in homeostasis and adaptation

IPC: C 07 K 14/34

**Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der
ursprünglichen Unterlagen dieser Patentanmeldung.**

München, den 1. Juli 2002
Deutsches Patent- und Markenamt

Der Präsident

Im Auftrag

A handwritten signature in black ink, appearing to read "F. K. K." or "F. K. K. K." followed by a stylized "K".

K. K. K.

CERTIFIED COPY OF
PRIORITY DOCUMENT

BEST AVAILABLE COPY

CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL PROTEINS

Background of the Invention

5 Certain products and by-products of naturally-occurring metabolic processes in cells have utility in a wide array of industries, including the food, feed, cosmetics, and pharmaceutical industries. These molecules, collectively termed 'fine chemicals', include organic acids, both proteinogenic and non-proteinogenic amino acids, nucleotides and nucleosides, lipids and fatty acids, diols, carbohydrates, aromatic 10 compounds, vitamins and cofactors, and enzymes. Their production is most conveniently performed through the large-scale culture of bacteria developed to produce and secrete large quantities of one or more desired molecules. One particularly useful organism for this purpose is *Corynebacterium glutamicum*, a gram positive, nonpathogenic bacterium. Through strain selection, a number of mutant strains have 15 been developed which produce an array of desirable compounds. However, selection of strains improved for the production of a particular molecule is a time-consuming and difficult process.

Summary of the Invention

20 This invention provides novel nucleic acid molecules which may be used to identify or classify *Corynebacterium glutamicum* or related species of bacteria. *C. glutamicum* is a gram positive, aerobic bacterium which is commonly used in industry for the large-scale production of a variety of fine chemicals, and also for the degradation of hydrocarbons (such as in petroleum spills) and for the oxidation of terpenoids. The 25 nucleic acid molecules therefore can be used to identify microorganisms which can be used to produce fine chemicals, e.g., by fermentation processes. While *C. glutamicum* itself is nonpathogenic, it is related to other *Corynebacterium* species, such as *Corynebacterium diphtheriae* (the causative agent of diphtheria), which are important human pathogens. The ability to identify the presence of *Corynebacterium* species therefore also can have significant clinical relevance, e.g., diagnostic applications. 30 Further, these nucleic acid molecules may serve as reference points for the mapping of the *C. glutamicum* genome, or of genomes of related organisms.

35 These novel nucleic acid molecules encode proteins, referred to herein as marker and fine chemical production (MCP) proteins. These MCP proteins may be involved, for example, in the direct or indirect production of one or more fine chemicals from *C. glutamicum*. The MCP proteins of the invention may also participate in the degradation of hydrocarbons or the oxidation of terpenoids. These proteins may also be utilized for

the identification of *Corynebacterium glutamicum* or organisms related to *C. glutamicum*: the presence of an MCP protein specific to *C. glutamicum* and related species in a mixture of proteins may indicate the presence of one of these bacteria in the sample. Further, these MCP proteins may have homologues in plants or animals which

5 are involved in a disease state or condition: these proteins thus may serve as useful pharmaceutical targets for drug screening and the development of therapeutic compounds.

Given the availability of cloning vectors for use in *Corynebacterium glutamicum*, such as those disclosed in Sinskey et al., U.S. Patent No. 4,649,119, and

10 techniques for genetic manipulation of *C. glutamicum* and the related *Brevibacterium* species (e.g., *lactofermenrum*) (Yoshihama et al. *J. Bacteriol.* 162: 591-597 (1985); Katsumata et al., *J. Bacteriol.* 159: 306-311 (1984); and Santamaria et al., *J. Gen Microbiol.* 130: 2237-2246 (1984)), the nucleic acid molecules of the invention may be utilized in the genetic engineering of this organism to modulate the production of one or

15 more fine chemicals. This modulation may be due to a direct effect of manipulation of a gene of the invention, or it may be due to an indirect effect of such manipulation. For example, by modifying the activity of a protein involved in the biosynthesis or degradation of a fine chemical (i.e., through mutagenesis of the corresponding gene), one may directly modulate the ability of the cell to synthesize or to degrade this

20 compound, thereby modulating the yield and/or efficiency of production of the fine chemical. Similarly, by modulating the activity of a protein which regulates a fine chemical metabolic pathway, one may directly influence whether the production of the desired compound is up- or down-regulated, either of which will modulate the yield or efficiency of production of the fine chemical from the cell.

25 Indirect modulation of fine chemical production may also result by modifying the activity of a protein of the invention (i.e., by mutagenesis of the corresponding gene) such that the overall ability of the cell to grow and divide or to remain viable and productive is increased. The production of fine chemicals from *C. glutamicum* is generally accomplished by the large-scale fermentative culture of these microorganisms.

30 conditions which are frequently suboptimal for growth and cell division. By engineering a protein of the invention (e.g., a stress response protein, a cell wall protein, or proteins involved in the metabolism of compounds necessary for cell growth and division to occur, such as nucleotides and amino acids) such that it is better able to survive, grow, and multiply in such conditions, it may be possible to increase the

35 number and productivity of such engineered *C. glutamicum* cells in large-scale culture, which in turn should result in increased yields and/or efficiency of production of one or more desired fine chemicals. Further, the metabolic pathways of any cell are necessarily

interrelated and coregulated. By altering the activity or regulation of any one metabolic pathway in *C. glutamicum* (i.e., by altering the activity of one of the proteins of the invention which participates in such a pathway), it is possible to concomitantly alter the activity or regulation of other metabolic pathways in this microorganism, which may be 5 directly involved in the synthesis or degradation of a fine chemical.

The invention provides novel nucleic acid molecules which encode proteins, referred to herein as MCP proteins, which are capable of, for example, modulating the production or efficiency of production of one or more fine chemicals from *C. glutamicum*, or of serving as identifying markers for *C. glutamicum* or related 10 organisms. Nucleic acid molecules encoding an MCP protein are referred to herein as MCP nucleic acid molecules. In a preferred embodiment, the MCP protein is capable of modulating the production or efficiency of production of one or more fine chemicals from *C. glutamicum*, or of serving as identifying markers for *C. glutamicum* or related organisms. Examples of such proteins include those encoded by the genes set forth in 15 Table 1.

Accordingly, one aspect of the invention pertains to isolated nucleic acid molecules (e.g., cDNAs) comprising a nucleotide sequence encoding an MCP protein or biologically active portions thereof, as well as nucleic acid fragments suitable as primers or hybridization probes for the detection or amplification of MCP-encoding nucleic acid 20 (e.g., DNA or mRNA). In particularly preferred embodiments, the isolated nucleic acid molecule comprises one of the nucleotide sequences set forth in Appendix A or the coding region or a complement thereof of one of these nucleotide sequences. In other particularly preferred embodiments, the isolated nucleic acid molecule of the invention comprises a nucleotide sequence which hybridizes to or is at least about 50%, preferably 25 at least about 60%, more preferably at least about 70%, 80% or 90%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to a nucleotide sequence set forth in Appendix A, or a portion thereof. In other preferred embodiments, the isolated nucleic acid molecule encodes one of the amino acid sequences set forth in Appendix B. The preferred MCP proteins of the present invention 30 also preferably possess at least one of the MCP activities described herein.

In another embodiment, the isolated nucleic acid molecule encodes a protein or portion thereof wherein the protein or portion thereof includes an amino acid sequence which is sufficiently homologous to an amino acid sequence of Appendix B, e.g., sufficiently homologous to an amino acid sequence of Appendix B such that the protein 35 or portion thereof maintains an MCP activity. Preferably, the protein or portion thereof encoded by the nucleic acid molecule maintains the ability to modulate the production or efficiency of production of one or more fine chemicals from *C. glutamicum*, or of

serving as an identifying marker for *C. glutamicum* or related organisms. In one embodiment, the protein encoded by the nucleic acid molecule is at least about 50%, preferably at least about 60%, and more preferably at least about 70%, 80%, or 90% and most preferably at least about 95%, 96%, 97%, 98%, or 99% or more homologous to an amino acid sequence of Appendix B (e.g., an entire amino acid sequence selected from those sequences set forth in Appendix B). In another preferred embodiment, the protein is a full length *C. glutamicum* protein which is substantially homologous to an entire amino acid sequence of Appendix B (encoded by an open reading frame shown in Appendix A).

10 In another preferred embodiment, the isolated nucleic acid molecule is derived from *C. glutamicum* and encodes a protein (e.g., an MCP fusion protein) which includes a biologically active domain which is at least about 50% or more homologous to one of the amino acid sequences of Appendix B and is able to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to 15 degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms, and which also includes heterologous nucleic acid sequences encoding a heterologous polypeptide or regulatory regions.

17 In another embodiment, the isolated nucleic acid molecule is at least 15 nucleotides in length and hybridizes under stringent conditions to a nucleic acid molecule comprising a nucleotide sequence of Appendix A. Preferably, the isolated nucleic acid molecule corresponds to a naturally-occurring nucleic acid molecule. More preferably, the isolated nucleic acid encodes a naturally-occurring *C. glutamicum* MCP protein, or a biologically active portion thereof.

22 Another aspect of the invention pertains to vectors, e.g., recombinant expression vectors, containing the nucleic acid molecules of the invention, and host cells into which such vectors have been introduced. In one embodiment, such a host cell is used to produce an MCP protein by culturing the host cell in a suitable medium. The MCP protein can then be isolated from the medium or the host cell.

27 Yet another aspect of the invention pertains to a genetically altered microorganism in which an MCP gene has been introduced or altered. In one embodiment, the genome of the microorganism has been altered by introduction of a nucleic acid molecule of the invention encoding wild-type or mutated MCP sequence as a transgene. In another embodiment, an endogenous MCP gene within the genome of the microorganism has been altered, e.g., functionally disrupted, by homologous recombination with an altered MCP gene. In a preferred embodiment, the microorganism belongs to the genus *Corynebacterium* or *Brevibacterium*, with

Corynebacterium glutamicum being particularly preferred. In a preferred embodiment, the microorganism is also utilized for the production of a desired compound, such as an amino acid, with lysine being particularly preferred.

Still another aspect of the invention pertains to an isolated MCP protein or a portion, e.g., a biologically active portion, thereof. In a preferred embodiment, the isolated MCP protein or portion thereof is capable of modulating the production or efficiency of production of one or more fine chemicals from *C. glutamicum*, or of serving as an identifying marker for *C. glutamicum* or related organisms. In another preferred embodiment, the isolated MCP protein or portion thereof is sufficiently homologous to an amino acid sequence of Appendix B such that the protein or portion thereof maintains the ability to, for example, modulate the production or efficiency of production of one or more fine chemicals from *C. glutamicum*, or to serve as identifying markers for *C. glutamicum* or related organisms.

The invention also provides an isolated preparation of an MCP protein. In preferred embodiments, the MCP protein comprises an amino acid sequence of Appendix B. In another preferred embodiment, the invention pertains to an isolated full length protein which is substantially homologous to an entire amino acid sequence of Appendix B (encoded by an open reading frame set forth in Appendix A). In yet another embodiment, the protein is at least about 50%, preferably at least about 60%, 20 and more preferably at least about 70%, 80%, or 90%, and most preferably at least about 95%, 96%, 97%, 98%, or 99% or more homologous to an entire amino acid sequence of Appendix B. In other embodiments, the isolated MCP protein comprises an amino acid sequence which is at least about 50% or more homologous to one of the amino acid sequences of Appendix B and is able to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms.

Alternatively, the isolated MCP protein can comprise an amino acid sequence which is encoded by a nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, or is at least about 50%, preferably at least about 60%, more preferably at least about 70%, 80%, or 90%, and even more preferably at least about 95%, 96%, 97%, 98%, or 99% or more homologous, to a nucleotide sequence of Appendix B. It is also preferred that the preferred forms of MCP proteins also have one or more of the MCP bioactivities described herein.

35 The MCP polypeptide, or a biologically active portion thereof, can be operatively linked to a non-MCP polypeptide to form a fusion protein. In preferred embodiments, this fusion protein has an activity which differs from that of the MCP

protein alone. In other preferred embodiments, this fusion protein is capable of modulating the yield, production and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, or of serving as an identifying marker for *C. glutamicum* or related organisms. In particularly preferred embodiments, integration of this fusion

5 protein into a host cell modulates production of a desired compound from the cell.

Another aspect of the invention pertains to a method for producing a fine chemical. This method involves the culturing of a cell containing a vector directing the expression of an MCP nucleic acid molecule of the invention, such that a fine chemical is produced. In a preferred embodiment, this method further includes the step of 10 obtaining a cell containing such a vector, in which a cell is transfected with a vector directing the expression of an MCP nucleic acid. In another preferred embodiment, this method further includes the step of recovering the fine chemical from the culture. In a particularly preferred embodiment, the cell is from the genus *Corynebacterium* or *Brevibacterium*, or is selected from those strains set forth in Table 3.

15 Another aspect of the invention pertains to methods for modulating production of a molecule from a microorganism. Such methods include contacting the cell with an agent which modulates MCP protein activity or MCP nucleic acid expression such that a cell associated activity is altered relative to this same activity in the absence of the agent. In a preferred embodiment, the cell is modulated for one or more *C. glutamicum* 20 MCP protein activities, such that the yield, production, and/or efficiency of production of a desired fine chemical by this microorganism is improved. The agent which modulates MCP protein activity can be an agent which stimulates MCP protein activity or MCP nucleic acid expression. Examples of agents which stimulate MCP protein activity or MCP nucleic acid expression include small molecules, active MCP proteins, and nucleic acids encoding MCP proteins that have been introduced into the cell. 25 Examples of agents which inhibit MCP activity or expression include small molecules and antisense MCP nucleic acid molecules.

Another aspect of the invention pertains to methods for modulating yields, production, and/or efficiency of production of a desired compound from a cell, 30 involving the introduction of a wild-type or mutant MCP gene into a cell, either maintained on a separate plasmid or integrated into the genome of the host cell. If integrated into the genome, such integration can be random, or it can take place by homologous recombination such that the native gene is replaced by the introduced copy, causing the production of the desired compound from the cell to be modulated. In a 35 preferred embodiment, said yields are increased. In another preferred embodiment, said chemical is a fine chemical. In a particularly preferred embodiment, said fine chemical is an amino acid. In especially preferred embodiments, said amino acid is L-lysine.

Detailed Description of the Invention

The present invention provides MCP nucleic acid and protein molecules. These MCP nucleic acid molecules may be utilized in the identification of *Corynebacterium glutamicum* or related organisms, in the mapping of the *C. glutamicum* genome (or a 5 genome of a closely related organism), or in the identification of microorganisms which may be used to produce fine chemicals, e.g., by fermentation processes. The proteins encoded by these nucleic acids may be utilized in the direct or indirect modulation of the production or efficiency of production of one or more fine chemicals from *C. glutamicum*, as identifying markers for *C. glutamicum* or related organisms, in the 10 oxidation of terpenoids or the degradation of hydrocarbons, or as targets for the development of therapeutic pharmaceutical compounds. Aspects of the invention are further explicated below.

I. Fine Chemicals

15 The term 'fine chemical' is art-recognized and includes molecules produced by an organism which have applications in various industries, such as, but not limited to, the pharmaceutical, agriculture, and cosmetics industries. Such compounds include organic acids, such as tartaric acid, itaconic acid, and diaminopimelic acid, both proteinogenic and non-proteinogenic amino acids, purine and pyrimidine bases, 20 nucleosides, and nucleotides (as described e.g. in Kuninaka, A. (1996) Nucleotides and related compounds, p. 561-612, in Biotechnology vol. 6, Rehm et al., eds. VCH: Weinheim, and references contained therein), lipids, both saturated and unsaturated fatty acids (e.g., arachidonic acid), diols (e.g., propane diol, and butane diol), carbohydrates (e.g., hyaluronic acid and trehalose), aromatic compounds (e.g., aromatic amines, 25 vanillin, and indigo), vitamins and cofactors (as described in Ullmann's Encyclopedia of Industrial Chemistry, vol. A27, "Vitamins", p. 443-613 (1996) VCH: Weinheim and references therein; and Ong, A.S., Niki, E. & Packer, L. (1995) "Nutrition, Lipids, Health, and Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia, and the Society for Free Radical Research - 30 Asia, held Sept. 1-3, 1994 at Penang, Malaysia, AOCS Press, (1995)). enzymes, and all other chemicals described in Gutcho (1983) Chemicals by Fermentation, Noyes Data Corporation, ISBN: 0818805086 and references therein. The metabolism and uses of certain of these fine chemicals are further explicated below.

35 A. Amino Acid Metabolism and Uses

Amino acids comprise the basic structural units of all proteins, and as such are essential for normal cellular functioning in all organisms. The term "amino acid" is art-

recognized. The proteinogenic amino acids, of which there are 20 species, serve as structural units for proteins, in which they are linked by peptide bonds, while the nonproteinogenic amino acids (hundreds of which are known) are not normally found in proteins (see Ullmann's Encyclopedia of Industrial Chemistry, vol. A2, p. 57-97 VCH: Weinheim (1985)). Amino acids may be in the D- or L- optical configuration, though L-amino acids are generally the only type found in naturally-occurring proteins.

5 Biosynthetic and degradative pathways of each of the 20 proteinogenic amino acids have been well characterized in both prokaryotic and eukaryotic cells (see, for example, Stryer, L. Biochemistry, 3rd edition, pages 578-590 (1988)). The 'essential' amino acids (histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophan, and valine), so named because they are generally a nutritional requirement due to the complexity of their biosynthesis, are readily converted by simple biosynthetic pathways to the remaining 11 'nonessential' amino acids (alanine, arginine, asparagine, aspartate, cysteine, glutamate, glutamine, glycine, proline, serine, and tyrosine). Higher animals

10 15 do retain the ability to synthesize some of these amino acids, but the essential amino acids must be supplied from the diet in order for normal protein synthesis to occur.

Aside from their function in protein biosynthesis, these amino acids are interesting chemicals in their own right, and many have been found to have various applications in the food, feed, chemical, cosmetics, agriculture, and pharmaceutical industries. Lysine is an important amino acid in the nutrition not only of humans, but also of monogastric animals such as poultry and swine. Glutamate is most commonly used as a flavor additive (mono-sodium glutamate, MSG) and is widely used throughout the food industry, as are aspartate, phenylalanine, glycine, and cysteine. Glycine, L-methionine and tryptophan are all utilized in the pharmaceutical industry. Glutamine, valine, leucine, isoleucine, histidine, arginine, proline, serine and alanine are of use in both the pharmaceutical and cosmetics industries. Threonine, tryptophan, and D/L-methionine are common feed additives. (Leuchtenberger, W. (1996) Amino acids - technical production and use, p. 466-502 in Rehm et al. (eds.) Biotechnology vol. 6, chapter 14a, VCH: Weinheim). Additionally, these amino acids have been found to be useful as precursors for the synthesis of synthetic amino acids and proteins, such as N-acetylcysteine, S-carboxymethyl-L-cysteine, (S)-5-hydroxytryptophan, and others described in Ullmann's Encyclopedia of Industrial Chemistry, vol. A2, p. 57-97, VCH: Weinheim, 1985.

The biosynthesis of these natural amino acids in organisms capable of producing them, such as bacteria, has been well characterized (for review of bacterial amino acid biosynthesis and regulation thereof, see Umbarger, H.E. (1978) *Ann. Rev. Biochem.* 47: 533-606). Glutamate is synthesized by the reductive amination of α -

ketoglutarate, an intermediate in the citric acid cycle. Glutamine, proline, and arginine are each subsequently produced from glutamate. The biosynthesis of serine is a three-step process beginning with 3-phosphoglycerate (an intermediate in glycolysis), and resulting in this amino acid after oxidation, transamination, and hydrolysis steps. Both 5 cysteine and glycine are produced from serine; the former by the condensation of homocysteine with serine, and the latter by the transferal of the side-chain β -carbon atom to tetrahydrofolate, in a reaction catalyzed by serine transhydroxymethylase. Phenylalanine, and tyrosine are synthesized from the glycolytic and pentose phosphate pathway precursors erythrose 4-phosphate and phosphoenolpyruvate in a 9-step 10 biosynthetic pathway that differ only at the final two steps after synthesis of prephenate. Tryptophan is also produced from these two initial molecules, but its synthesis is an 11-step pathway. Tyrosine may also be synthesized from phenylalanine, in a reaction catalyzed by phenylalanine hydroxylase. Alanine, valine, and leucine are all 15 biosynthetic products of pyruvate, the final product of glycolysis. Aspartate is formed from oxaloacetate, an intermediate of the citric acid cycle. Asparagine, methionine, threonine, and lysine are each produced by the conversion of aspartate. Isoleucine is formed from threonine. A complex 9-step pathway results in the production of histidine from 5-phosphoribosyl-1-pyrophosphate, an activated sugar.

Amino acids in excess of the protein synthesis needs of the cell cannot be stored. 20 and are instead degraded to provide intermediates for the major metabolic pathways of the cell (for review see Stryer, L. Biochemistry 3rd ed. Ch. 21 "Amino Acid Degradation and the Urea Cycle" p. 495-516 (1988)). Although the cell is able to convert unwanted amino acids into useful metabolic intermediates, amino acid production is costly in terms of energy, precursor molecules, and the enzymes necessary to synthesize them. 25 Thus it is not surprising that amino acid biosynthesis is regulated by feedback inhibition, in which the presence of a particular amino acid serves to slow or entirely stop its own production (for overview of feedback mechanisms in amino acid biosynthetic pathways, see Stryer, L. Biochemistry, 3rd ed. Ch. 24: "Biosynthesis of Amino Acids and Heme" p. 575-600 (1988)). Thus, the output of any particular amino acid is limited by the amount 30 of that amino acid present in the cell.

B. Vitamin, Cofactor, and Nutraceutical Metabolism and Uses

Vitamins, cofactors, and nutraceuticals comprise another group of molecules which the higher animals have lost the ability to synthesize and so must ingest, although 35 they are readily synthesized by other organisms such as bacteria. These molecules are either bioactive substances themselves, or are precursors of biologically active substances which may serve as electron carriers or intermediates in a variety of

metabolic pathways. Aside from their nutritive value, these compounds also have significant industrial value as coloring agents, antioxidants, and catalysts or other processing aids. (For an overview of the structure, activity, and industrial applications of these compounds, see, for example, Ullman's Encyclopedia of Industrial Chemistry, 5 "Vitamins" vol. A27, p. 443-613. VCH: Weinheim, 1996.) The term "vitamin" is art-
recognized, and includes nutrients which are required by an organism for normal functioning, but which that organism cannot synthesize by itself. The group of vitamins may encompass cofactors and nutraceutical compounds. The language "cofactor" includes nonproteinaceous compounds required for a normal enzymatic activity to occur. Such compounds may be organic or inorganic: the cofactor molecules of the invention are preferably organic. The term "nutraceutical" includes dietary supplements having health benefits in plants and animals, particularly humans. Examples of such molecules are vitamins, antioxidants, and also certain lipids (e.g., polyunsaturated fatty acids).

15 The biosynthesis of these molecules in organisms capable of producing them, such as bacteria, has been largely characterized (Ullman's Encyclopedia of Industrial Chemistry, "Vitamins" vol. A27, p. 443-613. VCH: Weinheim, 1996; Michal, G. (1999) Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology, John Wiley & Sons; Ong, A.S., Niki, E. & Packer, L. (1995) "Nutrition, Lipids, Health, and 20 Disease" Proceedings of the UNESCO/Confederation of Scientific and Technological Associations in Malaysia, and the Society for Free Radical Research - Asia, held Sept. 1-3, 1994 at Penang, Malaysia. AOCS Press: Champaign, IL X, 374 S).

Thiamin (vitamin B₁) is produced by the chemical coupling of pyrimidine and thiazole moieties. Riboflavin (vitamin B₂) is synthesized from guanosine-5'-triphosphate (GTP) and ribose-5'-phosphate. Riboflavin, in turn, is utilized for the synthesis of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD). The family of compounds collectively termed 'vitamin B₆' (e.g., pyridoxine, pyridoxamine, pyridoxal-5'-phosphate, and the commercially used pyridoxin hydrochloride) are all derivatives of the common structural unit, 5-hydroxy-6-methylpyridine. Pantothenate (pantothenic acid, (R)-(+)-N-(2,4-dihydroxy-3,3-dimethyl-1-oxobutyl)-β-alanine) can be produced either by chemical synthesis or by fermentation. The final steps in pantothenate biosynthesis consist of the ATP-driven condensation of β-alanine and pantoic acid. The enzymes responsible for the biosynthesis steps for the conversion to pantoic acid, to β-alanine and for the condensation to pantothenic acid are known. The metabolically active form of pantothenate is Coenzyme A, for which the biosynthesis proceeds in 5 enzymatic steps. Pantothenate, pyridoxal-5'-phosphate, cysteine and ATP are the precursors of Coenzyme A. These enzymes not only catalyze the formation of

panthothenate, but also the production of (R)-pantoic acid, (R)-pantolacton, (R)-panthenol (provitamin B₅), pantetheine (and its derivatives) and coenzyme A.

Biotin biosynthesis from the precursor molecule pimeloyl-CoA in

microorganisms has been studied in detail and several of the genes involved have been

5 identified. Many of the corresponding proteins have been found to also be involved in Fe-cluster synthesis and are members of the nifS class of proteins. Lipoic acid is derived from octanoic acid, and serves as a coenzyme in energy metabolism, where it becomes part of the pyruvate dehydrogenase complex and the α -ketoglutarate dehydrogenase complex. The folates are a group of substances which are all derivatives

10 of folic acid, which is turn is derived from L-glutamic acid, p-amino-benzoic acid and 6-methylpterin. The biosynthesis of folic acid and its derivatives, starting from the metabolism intermediates guanosine-5'-triphosphate (GTP), L-glutamic acid and p-amino-benzoic acid has been studied in detail in certain microorganisms.

Corrinoids (such as the cobalamines and particularly vitamin B₁₂) and

15 porphyrines belong to a group of chemicals characterized by a tetrapyrrole ring system. The biosynthesis of vitamin B₁₂ is sufficiently complex that it has not yet been completely characterized, but many of the enzymes and substrates involved are now known. Nicotinic acid (nicotinate), and nicotinamide are pyridine derivatives which are also termed 'niacin'. Niacin is the precursor of the important coenzymes NAD

20 (nicotinamide adenine dinucleotide) and NADP (nicotinamide adenine dinucleotide phosphate) and their reduced forms.

The large-scale production of these compounds has largely relied on cell-free chemical syntheses, though some of these chemicals have also been produced by large-scale culture of microorganisms, such as riboflavin, Vitamin B₆, pantothenate, and

25 biotin. Only Vitamin B₁₂ is produced solely by fermentation, due to the complexity of its synthesis. *In vitro* methodologies require significant inputs of materials and time, often at great cost.

C. Purine, Pyrimidine, Nucleoside and Nucleotide Metabolism and Uses

30 Purine and pyrimidine metabolism genes and their corresponding proteins are important targets for the therapy of tumor diseases and viral infections. The language "purine" or "pyrimidine" includes the nitrogenous bases which are constituents of nucleic acids, co-enzymes, and nucleotides. The term "nucleotide" includes the basic structural units of nucleic acid molecules, which are comprised of a nitrogenous base, a

35 pentose sugar (in the case of RNA, the sugar is ribose; in the case of DNA, the sugar is D-deoxyribose), and phosphoric acid. The language "nucleoside" includes molecules which serve as precursors to nucleotides, but which are lacking the phosphoric acid

moiety that nucleotides possess. By inhibiting the biosynthesis of these molecules, or their mobilization to form nucleic acid molecules, it is possible to inhibit RNA and DNA synthesis: by inhibiting this activity in a fashion targeted to cancerous cells, the ability of tumor cells to divide and replicate may be inhibited. Additionally, there are

5 nucleotides which may serve as energy stores (e.g., ADP, ATP) or as coenzymes (i.e., FAD and NAD).

Several publications have described the use of these chemicals for these medical indications, by influencing purine and/or pyrimidine metabolism (e.g. Christopherson, R.I. and Lyons, S.D. (1990) "Potent inhibitors of *de novo* pyrimidine and purine 10 biosynthesis as chemotherapeutic agents." *Med Res Reviews* 10: 505-548). Studies of enzymes involved in purine and pyrimidine metabolism have been focused on the development of new drugs which can be used, for example, as immunosuppressants or anti-proliferants (Smith, J.L. (1995) "Enzymes in nucleotide synthesis." *Curr. Opin. Struct. Biol.* 5: 752-757; (1995) *Biochem Soc. Transact.* 23: 877-902). However, purine 15 and pyrimidine bases, nucleosides and nucleotides have other utilities: as intermediates in the biosynthesis of several fine chemicals (e.g., thiamine, S-adenosyl-methionine, folates, or riboflavin), as energy carriers for the cell (e.g., ATP or GTP), and for chemicals themselves, commonly used as flavor enhancers (e.g., IMP or GMP) or for several medicinal applications (see, for example, Kuninaka, A. (1996) *Nucleotides and 20 Related Compounds in Biotechnology* vol. 6. Rehm et al., eds. VCH: Weinheim, p. 561-612). Also, enzymes involved in purine, pyrimidine, nucleoside, or nucleotide metabolism are increasingly serving as targets against which chemicals for crop protection, including fungicides, herbicides and insecticides, are developed.

The metabolism of these compounds in bacteria has been characterized (for 25 reviews see, for example, Zalkin, H. and Dixon, J.E. (1992) "*de novo* purine nucleotide biosynthesis", in: *Progress in Nucleic Acid Research and Molecular Biology*, vol. 42, Academic Press, p. 259-287; and Michal, G. (1999) "Nucleotides and Nucleosides". Chapter 8 in: *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, Wiley: New York). Purine metabolism has been the subject of intensive research, and is 30 essential to the normal functioning of the cell. Impaired purine metabolism in higher animals can cause severe disease, such as gout. Purine nucleotides are synthesized from ribose-5-phosphate, in a series of steps through the intermediate compound inosine-5'-phosphate (IMP), resulting in the production of guanosine-5'-monophosphate (GMP) or adenine-5'-monophosphate (AMP), from which the triphosphate forms utilized as 35 nucleotides are readily formed. These compounds are also utilized as energy stores, so their degradation provides energy for many different biochemical processes in the cell. Pyrimidine biosynthesis proceeds by the formation of uridine-5'-monophosphate (UMP)

from ribose-5-phosphate. UMP, in turn, is converted to cytidine-5'-triphosphate (CTP). The deoxy- forms of all of these nucleotides are produced in a one step reduction reaction from the diphosphate ribose form of the nucleotide to the diphosphate deoxyribose form of the nucleotide. Upon phosphorylation, these molecules are able to 5 participate in DNA synthesis.

D. Trehalose Metabolism and Uses

Trehalose consists of two glucose molecules, bound in a, α -1,1 linkage. It is commonly used in the food industry as a sweetener, an additive for dried or frozen 10 foods, and in beverages. However, it also has applications in the pharmaceutical, cosmetics and biotechnology industries (see, for example, Nishimoto et al., (1998) U.S. Patent No. 5,759,610; Singer, M.A. and Lindquist, S. (1998) *Trends Biotech.* 16: 460-467; Paiva, C.L.A. and Panek, A.D. (1996) *Bioitech. Ann. Rev.* 2: 293-314; and 15 Shiosaka, M. (1997) *J. Japan* 172: 97-102). Trehalose is produced by enzymes from many microorganisms and is naturally released into the surrounding medium, from which it can be collected using methods known in the art.

II. Elements and Methods of the Invention

The present invention is based, at least in part, on the discovery of novel 20 molecules, referred to herein as MCP nucleic acid molecules. These MCP nucleic acid molecules are useful not only for the identification of *C. glutamicum* or related bacterial species, but also as markers for the mapping of the *C. glutamicum* genome and in the identification of bacteria useful for the production of fine chemicals by, e.g., fermentative processes. The present invention is also based, at least in part, on the MCP 25 protein molecules encoded by these MCP nucleic acid molecules. These MCP proteins are capable of modulating the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, of serving as identifying markers for *C. glutamicum* or related organisms, of degrading hydrocarbons, and of serving as targets for the development of therapeutic pharmaceutical compounds. In one embodiment, the 30 MCP molecules of the invention directly or indirectly participate in one or more fine chemical metabolic pathways in *C. glutamicum*. In a preferred embodiment, the activity of the MCP molecules of the invention to indirectly or directly participate in such metabolic pathways has an impact on the production of a desired fine chemical by this microorganism. In a particularly preferred embodiment, the MCP molecules of the 35 invention are modulated in activity, such that the *C. glutamicum* metabolic pathways in which the MCP proteins of the invention participate are modulated in efficiency or

output, which either directly or indirectly modulates the production or efficiency of production of a desired fine chemical by *C. glutamicum*.

The language, "MCP protein" or "MCP polypeptide" includes proteins which are able to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target protein for drug screening or design, or to serve as identifying markers for *C. glutamicum* or related organisms. Examples of MCP proteins include those encoded by the MCP genes set forth in Table 1 and Appendix A. The terms "MCP gene" or "MCP nucleic acid sequence" include nucleic acid sequences encoding an MCP protein, which consist of a coding region and also corresponding untranslated 5' and 3' sequence regions. Examples of MCP genes include those set forth in Table 1. The terms "production" or "productivity" are art-recognized and include the concentration of the fermentation product (for example, the desired fine chemical) formed within a given time and a given fermentation volume (e.g., kg product per hour per liter). The term "efficiency of production" includes the time required for a particular level of production to be achieved (for example, how long it takes for the cell to attain a particular rate of output of a fine chemical). The term "yield" or "product/carbon yield" is art-recognized and includes the efficiency of the conversion of the carbon source into the product (i.e., fine chemical). This is generally written as, for example, kg product per kg carbon source. By increasing the yield or production of the compound, the quantity of recovered molecules, or of useful recovered molecules of that compound in a given amount of culture over a given amount of time is increased. The terms "biosynthesis" or a "biosynthetic pathway" are art-recognized and include the synthesis of a compound, preferably an organic compound, by a cell from intermediate compounds in what may be a multistep and highly regulated process. The terms "degradation" or a "degradation pathway" are art-recognized and include the breakdown of a compound, preferably an organic compound, by a cell to degradation products (generally speaking, smaller or less complex molecules) in what may be a multistep and highly regulated process. The language "metabolism" is art-recognized and includes the totality of the biochemical reactions that take place in an organism. The metabolism of a particular compound, then, (e.g., the metabolism of an amino acid such as glycine) comprises the overall biosynthetic, modification, and degradation pathways in the cell related to this compound.

In another embodiment, the MCP molecules of the invention are capable of modulating the production of a desired molecule, such as a fine chemical, in a microorganism such as *C. glutamicum*, either directly or indirectly. Using recombinant genetic techniques, one or more of the MCP proteins of the invention may be

manipulated such that its function is modulated. Such modulation of function may result in the modulation of the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*.

For example, by modifying the activity of a protein involved in the biosynthesis or degradation of a fine chemical (i.e., through mutagenesis of the corresponding gene), one may directly modulate the ability of the cell to synthesize or to degrade this compound, thereby modulating the yield and/or efficiency of production of the fine chemical. Similarly, by modulating the activity of a protein which regulates a fine chemical metabolic pathway, one may directly influence whether the production of the desired compound is up- or down-regulated, either of which will modulate the yield or efficiency of production of the fine chemical from the cell.

Indirect modulation of fine chemical production may also result by modifying the activity of a protein of the invention (i.e., by mutagenesis of the corresponding gene) such that the overall ability of the cell to grow and divide or to remain viable and productive is increased. The production of fine chemicals from *C. glutamicum* is generally accomplished by the large-scale fermentative culture of these microorganisms, conditions which are frequently suboptimal for growth and cell division. By engineering a protein of the invention (e.g., a stress response protein, a cell wall protein, or proteins involved in the metabolism of compounds necessary for cell growth and division to occur, such as nucleotides and amino acids) such that it is better able to survive, grow, and multiply in such conditions, it may be possible to increase the number and productivity of such engineered *C. glutamicum* cells in large-scale culture, which in turn should result in increased yields and/or efficiency of production of one or more desired fine chemicals. Further, the metabolic pathways of any cell are necessarily interrelated and coregulated. By altering the activity or regulation of any one metabolic pathway in *C. glutamicum* (i.e., by altering the activity of one of the proteins of the invention which participates in such a pathway), it is possible to concomitantly alter the activity or regulation of other metabolic pathways in this microorganism, which may be directly involved in the synthesis or degradation of a fine chemical.

The isolated nucleic acid sequences of the invention are contained within the genome of a *Corynebacterium glutamicum* strain available through the American Type Culture Collection, given designation ATCC 13032. The nucleotide sequences of the isolated *C. glutamicum* MCP nucleic acid molecules and the predicted amino acid sequences of the *C. glutamicum* MCP proteins are shown in Appendices A and B, respectively. Computational analyses were performed which classified and/or identified many of these nucleotide sequences as sequences having homology to *E. coli* or *Bacillus subtilis* genes.

The present invention also pertains to proteins which have an amino acid sequence which is substantially homologous to an amino acid sequence of Appendix B. As used herein, a protein which has an amino acid sequence which is substantially homologous to a selected amino acid sequence is least about 50% homologous to the 5 selected amino acid sequence, e.g., the entire selected amino acid sequence. A protein which has an amino acid sequence which is substantially homologous to a selected amino acid sequence can also be least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80%, 80-90%, or 90-95%, and most preferably at least about 96%, 97%, 98%, 99% or more homologous to the selected amino acid 10 sequence.

The MCP protein or a biologically active portion or fragment thereof of the invention is able to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying 15 marker for *C. glutamicum* or related organisms.

Various aspects of the invention are described in further detail in the following subsections:

A. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that 20 encode MCP polypeptides or biologically active portions thereof, as well as nucleic acid fragments sufficient for use as hybridization probes or primers for the identification or amplification of MCP-encoding nucleic acid (e.g., MCP DNA). These nucleic acid molecules may be used to identify *C. glutamicum* or related organisms, to map the genome of *C. glutamicum* or closely related bacteria, or to identify microorganisms 25 useful for the production of fine chemicals, e.g., by fermentative processes. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. This term also encompasses untranslated 30 sequence located at both the 3' and 5' ends of the coding region of the gene: at least about 100 nucleotides of sequence upstream from the 5' end of the coding region and at least about 20 nucleotides of sequence downstream from the 3' end of the coding region of the gene. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA. An "isolated" nucleic acid molecule is one which is 35 separated from other nucleic acid molecules which are present in the natural source of the nucleic acid. Preferably, an "isolated" nucleic acid is free of sequences which naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the

nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated MCP nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb or 0.1 kb of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived (e.g. a *C. glutamicum* cell). Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, e.g., a nucleic acid molecule having a nucleotide sequence of Appendix A, or a portion thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. For example, a *C. glutamicum* MCP cDNA can be isolated from a *C. glutamicum* library using all or portion of one of the sequences of Appendix A as a hybridization probe and standard hybridization techniques (e.g., as described in Sambrook, J., Fritsh, E. F., and 15 Maniatis, T. *Molecular Cloning. A Laboratory Manual*. 2nd. ed. Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989). Moreover, a nucleic acid molecule encompassing all or a portion of one of the sequences of Appendix A can be isolated by the polymerase chain reaction using oligonucleotide primers designed based upon this sequence (e.g., a nucleic acid molecule encompassing 20 all or a portion of one of the sequences of Appendix A can be isolated by the polymerase chain reaction using oligonucleotide primers designed based upon this same sequence of Appendix A). For example, mRNA can be isolated from normal endothelial cells (e.g., by the guanidinium-thiocyanate extraction procedure of Chirgwin et al. (1979) *Biochemistry* 18: 5294-5299) and cDNA can be prepared using reverse transcriptase 25 (e.g., Moloney MLV reverse transcriptase, available from Gibco/BRL, Bethesda, MD; or AMV reverse transcriptase, available from Seikagaku America, Inc., St. Petersburg, FL) and random polynucleotide primers or oligonucleotide primers based upon one of the nucleotide sequences shown in Appendix A. Synthetic oligonucleotide primers for polymerase chain reaction amplification can be designed based upon one of the 30 nucleotide sequences shown in Appendix A. A nucleic acid of the invention can be amplified using cDNA or, alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to an MCP 35 nucleotide sequence can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

In a preferred embodiment, an isolated nucleic acid molecule of the invention comprises one of the nucleotide sequences shown in Appendix A. The sequences of Appendix A correspond to the *Corynebacterium glutamicum* MCP cDNAs of the invention. This cDNA comprises sequences encoding MCP proteins (i.e., "the coding region", indicated in each sequence in Appendix A), as well as 5' untranslated sequences and 3' untranslated sequences, also indicated in Appendix A. Alternatively, the nucleic acid molecule can comprise only the coding region of any of the sequences in Appendix A.

For the purposes of this application, it will be understood that each of the sequences set forth in Appendix A has an identifying RXA number having the designation "RXA" followed by 5 digits (i.e., RXA00003). Each of these sequences comprises up to three parts: a 5' upstream region, a coding region, and a downstream region. Each of these three regions is identified by the same RXA designation to eliminate confusion. The recitation "one of the sequences in Appendix A", then, refers to any of the sequences in Appendix A, which may be distinguished by their differing RXA designations. The coding region of each of these sequences is translated into a corresponding amino acid sequence, which is set forth in Appendix B. The sequences of Appendix B are identified by the same RXA designations as Appendix A, such that they can be readily correlated. For example, the amino acid sequence in Appendix B designated RXA00003 is a translation of the coding region of the nucleotide sequence of nucleic acid molecule RXA00003 in Appendix A.

In one embodiment, the nucleic acid molecules of the present invention are not intended to include those compiled in Table 2.

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of one of the nucleotide sequences shown in Appendix A, or a portion thereof. A nucleic acid molecule which is complementary to one of the nucleotide sequences shown in Appendix A is one which is sufficiently complementary to one of the nucleotide sequences shown in Appendix A such that it can hybridize to one of the nucleotide sequences shown in Appendix A, thereby forming a stable duplex.

In still another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleotide sequence which is at least about 50-60%, preferably at least about 60-70%, more preferably at least about 70-80%, 80-90%, or 90-95%, and even more preferably at least about 95%, 96%, 97%, 98%, 99% or more homologous to a nucleotide sequence shown in Appendix A, or a portion thereof. In an additional preferred embodiment, an isolated nucleic acid molecule of the invention comprises a

nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, to one of the nucleotide sequences shown in Appendix A; or a portion thereof.

Moreover, the nucleic acid molecule of the invention can comprise only a portion of the coding region of one of the sequences in Appendix A, for example a fragment which can be used as a probe or primer or a fragment encoding a biologically active portion of an MCP protein. The nucleotide sequences determined from the cloning of the MCP genes from *C. glutamicum* allows for the generation of probes and primers designed for use in identifying and/or cloning MCP homologues in other cell types and organisms, as well as MCP homologues from other *Corynebacteria* or related species. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, preferably about 25, more preferably about 40, 50 or 75 consecutive nucleotides of a sense strand of one of the sequences set forth in Appendix A, an anti-sense sequence of one of the sequences set forth in Appendix A, or naturally occurring mutants thereof. Primers based on a nucleotide sequence of Appendix A can be used in PCR reactions to clone MCP homologues. Probes based on the MCP nucleotide sequences can be used to detect transcripts or genomic sequences encoding the same or homologous proteins. In preferred embodiments, the probe further comprises a label group attached thereto, e.g. the label group can be a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as a part of a diagnostic test kit for identifying cells which misexpress an MCP protein, such as by measuring a level of an MCP-encoding nucleic acid in a sample of cells, e.g., detecting MCP mRNA levels or determining whether a genomic MCP gene has been mutated or deleted.

In one embodiment, the nucleic acid molecule of the invention encodes a protein or portion thereof which includes an amino acid sequence which is sufficiently homologous to an amino acid sequence of Appendix B such that the protein or portion thereof maintains the ability to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms. As used herein, the language "sufficiently homologous" refers to proteins or portions thereof which have amino acid sequences which include a minimum number of identical or equivalent (e.g., an amino acid residue which has a similar side chain as an amino acid residue in one of the sequences of Appendix B) amino acid residues to an amino acid sequence of Appendix B such that the protein or portion thereof is able to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C.*

glutamicum, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms. Examples of such activities are also described herein. Thus, "the function of an MCP protein" contributes to the overall regulation of one or more fine chemical 5 metabolic pathways, or to the degradation of a hydrocarbon, or to the oxidation of a terpenoid.

In another embodiment, the protein is at least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80%, 80-90%, 90-95%, and most preferably at least about 96%, 97%, 98%, 99% or more homologous to an entire amino 10 acid sequence of Appendix B.

Portions of proteins encoded by the MCP nucleic acid molecules of the invention are preferably biologically active portions of one of the MCP proteins. As used herein, the term "biologically active portion of an MCP protein" is intended to include a portion, e.g., a domain/motif, of an MCP protein that modulates the yield, production, and/or 15 efficiency of production of one or more fine chemicals from *C. glutamicum*, that degrades hydrocarbons, that oxidizes terpenoids, that may serve as a target for drug development, or that may serve as an identifying marker for *C. glutamicum* or related organisms. To determine whether an MCP protein or a biologically active portion thereof can modulate the yield, production, and/or efficiency of production of one or 20 more fine chemicals from *C. glutamicum*, can degrade hydrocarbons, or can oxidize terpenoids, an assay of activity may be performed. Such assay methods are well known to those skilled in the art, as detailed in Example 8 of the Exemplification.

Additional nucleic acid fragments encoding biologically active portions of an MCP protein can be prepared by isolating a portion of one of the sequences in Appendix 25 B, expressing the encoded portion of the MCP protein or peptide (e.g., by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the MCP protein or peptide.

The invention further encompasses nucleic acid molecules that differ from one of the nucleotide sequences shown in Appendix A (and portions thereof) due to degeneracy 30 of the genetic code and thus encode the same MCP protein as that encoded by the nucleotide sequences shown in Appendix A. In another embodiment, an isolated nucleic acid molecule of the invention has a nucleotide sequence encoding a protein having an amino acid sequence shown in Appendix B. In a still further embodiment, the nucleic acid molecule of the invention encodes a full length *C. glutamicum* protein which is 35 substantially homologous to an amino acid sequence of Appendix B (encoded by an open reading frame shown in Appendix A).

In addition to the *C. glutamicum* MCP nucleotide sequences shown in Appendix A, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequences of MCP proteins may exist within a population (e.g., the *C. glutamicum* population). Such genetic polymorphism in the 5 MCP gene may exist among individuals within a population due to natural variation. As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding an MCP protein, preferably a *C. glutamicum* MCP protein. Such natural variations can typically result in 1-5% variance in the nucleotide sequence of the MCP gene. Any and all such nucleotide variations and 10 resulting amino acid polymorphisms in MCP that are the result of natural variation and that do not alter the functional activity of MCP proteins are intended to be within the scope of the invention.

Nucleic acid molecules corresponding to natural variants and non-*C. glutamicum* homologues of the *C. glutamicum* MCP cDNA of the invention can be isolated based on 15 their homology to the *C. glutamicum* MCP nucleic acid disclosed herein using the *C. glutamicum* cDNA, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 15 nucleotides in length and hybridizes under stringent conditions to the nucleic acid 20 molecule comprising a nucleotide sequence of Appendix A. In other embodiments, the nucleic acid is at least 30, 50, 100, 250 or more nucleotides in length. As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% homologous to each other typically remain hybridized to each other. Preferably, the conditions are 25 such that sequences at least about 65%, more preferably at least about 70%, and even more preferably at least about 75% or more homologous to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A preferred, non-limiting example of stringent hybridization 30 conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C. followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C. Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to a sequence of Appendix A corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring" nucleic acid molecule refers to an 35 RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein). In one embodiment, the nucleic acid encodes a natural *C. glutamicum* MCP protein.

In addition to naturally-occurring variants of the MCP sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation into a nucleotide sequence of Appendix A, thereby leading to changes in the amino acid sequence of the encoded MCP protein, without altering the functional ability 5 of the MCP protein. For example, nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues can be made in a sequence of Appendix A. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence of one of the MCP proteins (Appendix B) without altering the activity of said MCP protein, whereas an "essential" amino acid residue is required for 10 MCP protein activity. Other amino acid residues, however, (e.g., those that are not conserved or only semi-conserved in the domain having MCP activity) may not be essential for activity and thus are likely to be amenable to alteration without altering MCP activity.

Accordingly, another aspect of the invention pertains to nucleic acid molecules 15 encoding MCP proteins that contain changes in amino acid residues that are not essential for MCP activity. Such MCP proteins differ in amino acid sequence from a sequence contained in Appendix B yet retain at least one of the MCP activities described herein. In one embodiment, the isolated nucleic acid molecule comprises a nucleotide sequence encoding a protein, wherein the protein comprises an amino acid sequence at least about 20 50% homologous to an amino acid sequence of Appendix B and is able to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms. Preferably, the protein encoded by the nucleic acid molecule is at least about 25 50-60% homologous to one of the sequences in Appendix B, more preferably at least about 60-70% homologous to one of the sequences in Appendix B, even more preferably at least about 70-80%, 80-90%, 90-95% homologous to one of the sequences in Appendix B, and most preferably at least about 96%, 97%, 98%, or 99% homologous to one of the sequences in Appendix B.

30 To determine the percent homology of two amino acid sequences (e.g., one of the sequences of Appendix B and a mutant form thereof) or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of one protein or nucleic acid for optimal alignment with the other protein or nucleic acid). The amino acid residues or nucleotides at corresponding amino acid 35 positions or nucleotide positions are then compared. When a position in one sequence (e.g., one of the sequences of Appendix B) is occupied by the same amino acid residue or nucleotide as the corresponding position in the other sequence (e.g., a mutant form of

the sequence selected from Appendix B), then the molecules are homologous at that position (i.e., as used herein amino acid or nucleic acid "homology" is equivalent to amino acid or nucleic acid "identity"). The percent homology between the two sequences is a function of the number of identical positions shared by the sequences

5 (i.e., % homology = # of identical positions/total # of positions x 100).

An isolated nucleic acid molecule encoding an MCP protein homologous to a protein sequence of Appendix B can be created by introducing one or more nucleotide substitutions, additions or deletions into a nucleotide sequence of Appendix A such that one or more amino acid substitutions, additions or deletions are introduced into the 10 encoded protein. Mutations can be introduced into one of the sequences of Appendix A by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a 15 similar side chain. Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, 20 proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Thus, a predicted nonessential amino acid residue in an MCP protein is preferably replaced with another amino acid residue from the same side chain 25 family. Alternatively, in another embodiment, mutations can be introduced randomly along all or part of an MCP coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for an MCP activity described herein to identify mutants that retain MCP activity. Following mutagenesis of one of the sequences of Appendix A, the encoded protein can be expressed recombinantly and the activity of the protein can be determined using, for example, assays described herein (see Example 8 of 30 the Exemplification).

In addition to the nucleic acid molecules encoding MCP proteins described above, another aspect of the invention pertains to isolated nucleic acid molecules which are antisense thereto. An "antisense" nucleic acid comprises a nucleotide sequence which is complementary to a "sense" nucleic acid encoding a protein, e.g., 35 complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be

complementary to an entire MCP coding strand, or to only a portion thereof. In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence encoding an MCP protein. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are

5 translated into amino acid residues (e.g., the entire coding region of SEQ ID RXA00003 comprises nucleotides 1 to 741). In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence encoding MCP. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred

10 to as 5' and 3' untranslated regions).

Given the coding strand sequences encoding MCP disclosed herein (e.g., the sequences set forth in Appendix A), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of MCP mRNA, but more preferably is an oligonucleotide which is antisense to only a portion of the coding or noncoding region of MCP mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of MCP mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed

15 by chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic

20 acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxymethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil,

25 dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiacytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-

amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

5 The antisense nucleic acid molecules of the invention are typically administered to a cell or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding an MCP protein to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by 10 conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. The antisense molecule can be modified such that it specifically binds to a receptor or an antigen expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecule to a peptide or 15 an antibody which binds to a cell surface receptor or antigen. The antisense nucleic acid molecule can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a eubacterial, viral or eucaryotic promoter are preferred.

20 In yet another embodiment, the antisense nucleic acid molecule of the invention is an α -anomeric nucleic acid molecule. An α -anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β -units, the strands run parallel to each other (Gaultier et al. (1987) *Nucleic Acids Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o- 25 methylribonucleotide (Inoue et al. (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue et al. (1987) *FEBS Lett.* 215:327-330).

30 In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave MCP mRNA transcripts to thereby inhibit translation of MCP mRNA. A ribozyme having specificity for an MCP-encoding nucleic acid can be designed based upon the nucleotide sequence of an MCP cDNA disclosed herein (i.e., 35 RXA00003 in Appendix A). For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in an MCP-encoding mRNA.

See, e.g., Cech et al. U.S. Patent No. 4,987,071 and Cech et al. U.S. Patent No. 5,116,742. Alternatively, MCP mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel, D. and Szostak, J.W. (1993) *Science* 261:1411-1418.

5 Alternatively, MCP gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region of an MCP nucleotide sequence (e.g., an MCP promoter and/or enhancers) to form triple helical structures that prevent transcription of an MCP gene in target cells. See generally, Helene, C. (1991) *Anticancer Drug Des.* 6(6):569-84; Helene, C. et al. (1992) *Ann. N.Y. Acad. Sci.* 660:27-10 36; and Maher, L.J. (1992) *Bioassays* 14(12):807-15.

B. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding an MCP protein (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of 15 autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to 20 which they are operatively linked. Such vectors are referred to herein as "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, 25 such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell, which means that the recombinant expression vectors include one or more regulatory 30 sequences, selected on the basis of the host cells to be used for expression, which is operatively linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of

interest is linked to the regulatory sequence(s) in a manner which allows for expression of the nucleotide sequence (e.g., in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, repressor binding sites, activator binding sites, enhancer regions and other expression control elements (e.g., terminators, other elements of mRNA secondary structure, or polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185. Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in 5 many types of host cell and those which direct expression of the nucleotide sequence only in certain host cells. It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides. 10 including fusion proteins or peptides, encoded by nucleic acids as described herein (e.g., MCP proteins, mutant forms of MCP proteins, fusion proteins, etc.).

The recombinant expression vectors of the invention can be designed for expression of MCP proteins in prokaryotic or eukaryotic cells. For example, MCP genes can be expressed in bacterial cells such as *C. glutamicum*, insect cells (using 15 baculovirus expression vectors), yeast and other fungal cells (see Romanos, M.A. et al. (1992) "Foreign gene expression in yeast: a review", *Yeast* 8: 423-488; van den Hondel, C.A.M.J.J. et al. (1991) "Heterologous gene expression in filamentous fungi" in: *More Gene Manipulations in Fungi*, J.W. Bennet & L.L. Lasure, eds., p. 396-428: Academic Press: San Diego; and van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) "Gene transfer 20 systems and vector development for filamentous fungi, in: *Applied Molecular Genetics of Fungi*, Peberdy, J.F. et al., eds., p. 1-28. Cambridge University Press: Cambridge). 25 algae and multicellular plant cells (see Schmidt, R. and Willmitzer, L. (1988) High efficiency *Agrobacterium tumefaciens* -mediated transformation of *Arabidopsis thaliana* leaf and cotyledon explants" *Plant Cell Rep.*: 583-586), or mammalian cells. 30 Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185. Academic Press, San Diego, CA (1990). Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out with vectors 35 containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion

vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase.

5 Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith, D.B. and Johnson, K.S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRJT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein. In one embodiment, the coding sequence of the MCP protein is cloned into a pGEX expression vector to create a vector encoding a fusion protein comprising, from the N-terminus to the C-terminus, GST-thrombin cleavage site-X protein. The fusion 10 protein can be purified by affinity chromatography using glutathione-agarose resin. Recombinant MCP protein unfused to GST can be recovered by cleavage of the fusion 15 protein with thrombin.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann et al., (1988) *Gene* 69:301-315) and pET 11d (Studier et al., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral 20 polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident λ prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression is to express the 30 protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those 35 preferentially utilized in the bacterium chosen for expression, such as *C. glutamicum* (Wada et al. (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the MCP protein expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerevisiae* include pYEpSec1 (Baldari, et al., (1987) *Embo J.* 6:229-234), pMFA (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz et al., (1987) *Gene* 54:113-123), and pYES2 (Invitrogen Corporation, San Diego, CA). Vectors and methods for the construction of vectors appropriate for use in other fungi, such as the filamentous fungi, include those detailed in: van den Hondel, C.A.M.J.J. & Punt, P.J. (1991) "Gene transfer systems and vector development for filamentous fungi", in: *Applied Molecular Genetics of Fungi*, J.F. Peberdy, et al., eds., p. 1-28, Cambridge University Press: Cambridge.

10 Alternatively, the MCP proteins of the invention can be expressed in insect cells using baculovirus expression vectors. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith et al. (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

15 In another embodiment, the MCP proteins of the invention may be expressed in unicellular plant cells (such as algae) or in plant cells from higher plants (e.g., the spermatophytes, such as crop plants). Examples of plant expression vectors include those detailed in: Becker, D., Kemper, E., Schell, J. and Masterson, R. (1992) "New plant binary vectors with selectable markers located proximal to the left border", *Plant Mol. Biol.* 20: 1195-1197; and Bevan, M.W. (1984) "Binary *Agrobacterium* vectors for plant transformation", *Nucl. Acid. Res.* 12: 8711-8721.

20 In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed, B. (1987) *Nature* 329:840) and pMT2PC (Kaufman et al. (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook, J., Fritsh, E. F., and Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2nd. ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

25 In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert et al.

(1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji et al. (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters 5 (e.g., the neurofilament promoter; Byrne and Ruddle (1989) *PNAS* 86:5473-5477), pancreas-specific promoters (Edlund et al. (1985) *Science* 230:912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) 10 *Science* 249:374-379) and the α -fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operatively linked to a regulatory sequence in 15 a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to MCP mRNA. Regulatory sequences operatively linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which 20 direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene 25 expression using antisense genes see Weintraub, H. et al. (1986) "Antisense RNA as a molecular tool for genetic analysis", *Reviews - Trends in Genetics*, Vol. 1(1).

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such 30 terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

35 A host cell can be any prokaryotic or eukaryotic cell. For example, an MCP protein can be expressed in bacterial cells such as *C. glutamicum*, insect cells, yeast or mammalian cells (such as Chinese hamster ovary cells (CHO) or COS cells). Other

suitable host cells are known to those skilled in the art. Microorganisms related to *Corynebacterium glutamicum* which may be conveniently used as host cells for the nucleic acid and protein molecules of the invention are set forth in Table 3.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via

- 5 conventional transformation or transfection techniques. As used herein, the terms "transformation", "transfection", "conjugation" and "transduction" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including using natural competence, chemical mediated transfer, calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated
- 10 transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, et al. (*Molecular Cloning. A Laboratory Manual*. 2nd. ed. Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the

- 15 expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker can be introduced into a host cell on the same vector as that encoding an MCP protein or can be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by, for example, drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).
- 20

25 To create a homologous recombinant microorganism, a vector is prepared which contains at least a portion of an MCP gene into which a deletion, addition or substitution has been introduced to thereby alter, e.g., functionally disrupt, the MCP gene. Preferably, this MCP gene is a *Corynebacterium glutamicum* MCP gene, but it can be a homologue from a related bacterium or even from a mammalian, yeast, or insect source.

- 30 In a preferred embodiment, the vector is designed such that, upon homologous recombination, the endogenous MCP gene is functionally disrupted (i.e., no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous MCP gene is mutated or otherwise altered but still encodes functional protein (e.g., the upstream regulatory region can be altered to thereby alter the expression of the endogenous MCP protein). In the homologous recombination vector, the altered portion of the MCP gene is flanked at its 5' and 3' ends by additional nucleic acid of the MCP
- 35

gene to allow for homologous recombination to occur between the exogenous MCP gene carried by the vector and an endogenous MCP gene in a microorganism. The additional flanking MCP nucleic acid is of sufficient length for successful homologous recombination with the endogenous gene. Typically, less than one kilobase of flanking

5 DNA (both at the 5' and 3' ends) is included in the vector (see e.g., Thomas, K.R. and Capecchi, M.R. (1987) *Cell* 51: 503 for a description of homologous recombination vectors). The vector is introduced into a microorganism (e.g., by electroporation) and cells in which the introduced MCP gene has homologously recombined with the endogenous MCP gene are selected, using art-known techniques.

10 In another embodiment, recombinant microorganisms can be produced which contain selected systems which allow for regulated expression of the introduced gene. For example, inclusion of an MCP gene on a vector placing it under control of the lac operon permits expression of the MCP gene in the presence of IPTG. Such regulatory systems are well known in the art.

15 A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce (i.e., express) an MCP protein. Accordingly, the invention further provides methods for producing MCP proteins using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding an MCP protein has been introduced, or into which genome has been introduced a gene encoding a wild-type or altered MCP protein) in a suitable medium until MCP protein is produced. In another embodiment, the method further comprises isolating MCP proteins from the medium or the host cell.

20

25 C. *Isolated MCP Proteins*

Another aspect of the invention pertains to isolated MCP proteins, and biologically active portions thereof. An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material when produced by recombinant DNA techniques, or chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of MCP protein in which the protein is separated from cellular components of the cells in which it is naturally or recombinantly produced. In one embodiment, the language "substantially free of cellular material" includes preparations of MCP protein having less than about 30% (by dry weight) of non-MCP protein (also referred to herein as a "contaminating protein"), more preferably less than about 20% of non-MCP protein, still more preferably less than about 10% of non-MCP protein, and most preferably less than about 5% non-MCP protein. When the MCP protein or biologically active portion

30

35

thereof is recombinantly produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, more preferably less than about 10%, and most preferably less than about 5% of the volume of the protein preparation. The language "substantially free of chemical precursors or other chemicals" includes preparations of MCP protein in which the protein is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein. In one embodiment, the language "substantially free of chemical precursors or other chemicals" includes preparations of MCP protein having less than about 30% (by dry weight) of chemical precursors or non-MCP chemicals, more preferably less than about 20% chemical precursors or non-MCP chemicals, still more preferably less than about 10% chemical precursors or non-MCP chemicals, and most preferably less than about 5% chemical precursors or non-MCP chemicals. In preferred embodiments, isolated proteins or biologically active portions thereof lack contaminating proteins from the same organism from which the MCP protein is derived. Typically, such proteins are produced by recombinant expression of, for example, a *C. glutamicum* MCP protein in a microorganism such as *C. glutamicum*.

An isolated MCP protein or a portion thereof of the invention is able to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms. In preferred embodiments, the protein or portion thereof comprises an amino acid sequence which is sufficiently homologous to an amino acid sequence of Appendix B such that the protein or portion thereof maintains the ability to modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms. The portion of the protein is preferably a biologically active portion as described herein. In another preferred embodiment, an MCP protein of the invention has an amino acid sequence shown in Appendix B. In yet another preferred embodiment, the MCP protein has an amino acid sequence which is encoded by a nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, to a nucleotide sequence of Appendix A. In still another preferred embodiment, the MCP protein has an amino acid sequence which is encoded by a nucleotide sequence that is at least about 50-60%, preferably at least about 60-70%, more preferably at least about 70-80%, 80-90%, 90-95%, and even more preferably at least about 96%, 97%, 98%, 99% or more homologous to one of the amino acid sequences of Appendix B. The preferred MCP proteins of the present invention also preferably possess at least one of the MCP

activities described herein. For example, a preferred MCP protein of the present invention includes an amino acid sequence encoded by a nucleotide sequence which hybridizes, e.g., hybridizes under stringent conditions, to a nucleotide sequence of Appendix A, and which is able to modulate the yield, production, and/or efficiency of 5 production of one or more fine chemicals from *C. glutamicum*, to degrade hydrocarbons, to oxidize terpenoids, to serve as a target for drug development, or to serve as an identifying marker for *C. glutamicum* or related organisms.

In other embodiments, the MCP protein is substantially homologous to an amino acid sequence of Appendix B and retains the functional activity of the protein of one of 10 the sequences of Appendix B yet differs in amino acid sequence due to natural variation or mutagenesis, as described in detail in subsection I above. Accordingly, in another embodiment, the MCP protein is a protein which comprises an amino acid sequence which is at least about 50-60%, preferably at least about 60-70%, and more preferably at least about 70-80, 80-90, 90-95%, and most preferably at least about 96%, 97%, 98%, 15 99% or more homologous to an entire amino acid sequence of Appendix B and which has at least one of the MCP activities described herein. In another embodiment, the invention pertains to a full length *C. glutamicum* protein which is substantially homologous to an entire amino acid sequence of Appendix B.

Biologically active portions of an MCP protein include peptides comprising 20 amino acid sequences derived from the amino acid sequence of an MCP protein, e.g., an amino acid sequence shown in Appendix B or the amino acid sequence of a protein homologous to an MCP protein, which include fewer amino acids than a full length MCP protein or the full length protein which is homologous to an MCP protein, and exhibit at least one activity of an MCP protein. Typically, biologically active portions 25 (peptides, e.g., peptides which are, for example, 5, 10, 15, 20, 30, 35, 36, 37, 38, 39, 40, 50, 100 or more amino acids in length) comprise a domain or motif with at least one activity of an MCP protein. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the activities described herein. Preferably, the biologically 30 active portions of an MCP protein include one or more selected domains/motifs or portions thereof having biological activity.

MCP proteins are preferably produced by recombinant DNA techniques. For example, a nucleic acid molecule encoding the protein is cloned into an expression vector (as described above), the expression vector is introduced into a host cell (as 35 described above) and the MCP protein is expressed in the host cell. The MCP protein can then be isolated from the cells by an appropriate purification scheme using standard protein purification techniques. Alternative to recombinant expression, an MCP protein,

polypeptide, or peptide can be synthesized chemically using standard peptide synthesis techniques. Moreover, native MCP protein can be isolated from cells (e.g., endothelial cells, bacterial cells, fungal cells or other cells), for example using an anti-MCP antibody, which can be produced by standard techniques utilizing an MCP protein or 5 fragment thereof of this invention.

The invention also provides MCP chimeric or fusion proteins. As used herein, an MCP "chimeric protein" or "fusion protein" comprises an MCP polypeptide operatively linked to a non-MCP polypeptide. An "MCP polypeptide" refers to a polypeptide having an amino acid sequence corresponding to an MCP protein, whereas a 10 "non-MCP polypeptide" refers to a polypeptide having an amino acid sequence corresponding to a protein which is not substantially homologous to the MCP protein, e.g., a protein which is different from the MCP protein and which is derived from the same or a different organism. Within the fusion protein, the term "operatively linked" is intended to indicate that the MCP polypeptide and the non-MCP polypeptide are fused 15 in-frame to each other. The non-MCP polypeptide can be fused to the N-terminus or C-terminus of the MCP polypeptide. For example, in one embodiment the fusion protein is a GST-MCP fusion protein in which the MCP sequences are fused to the C-terminus of the GST sequences. Such fusion proteins can facilitate the purification of recombinant MCP proteins. In another embodiment, the fusion protein is an MCP 20 protein containing a heterologous signal sequence at its N-terminus. In certain host cells (e.g., mammalian host cells, bacterial host cells, fungal host cells), expression and/or secretion of an MCP protein can be increased through use of a heterologous signal sequence.

Preferably, an MCP chimeric or fusion protein of the invention is produced by 25 standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, for example by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid 30 undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene 35 fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for example, *Current Protocols in Molecular Biology*, eds. Ausubel et al. John Wiley & Sons: 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). An MCP-

encoding nucleic acid can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the MCP protein.

Homologues of the MCP protein can be generated by mutagenesis, e.g., discrete point mutation or truncation of the MCP protein. As used herein, the term "homologue" 5 refers to a variant form of the MCP protein which acts as an agonist or antagonist of the activity of the MCP protein. An agonist of the MCP protein can retain substantially the same, or a subset, of the biological activities of the MCP protein. An antagonist of the MCP protein can inhibit one or more of the activities of the naturally occurring form of the MCP protein, by, for example, competitively binding to a downstream or upstream 10 member of a biochemical pathway which includes the MCP protein.

In an alternative embodiment, homologues of the MCP protein can be identified by screening combinatorial libraries of mutants, e.g., truncation mutants, of the MCP protein for MCP protein agonist or antagonist activity. In one embodiment, a variegated library of MCP variants is generated by combinatorial mutagenesis at the nucleic acid 15 level and is encoded by a variegated gene library. A variegated library of MCP variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential MCP sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display) containing the set of MCP sequences therein. 20 There are a variety of methods which can be used to produce libraries of potential MCP homologues from a degenerate oligonucleotide sequence. Chemical synthesis of a degenerate gene sequence can be performed in an automatic DNA synthesizer, and the synthetic gene then ligated into an appropriate expression vector. Use of a degenerate set of genes allows for the provision, in one mixture, of all of the sequences encoding 25 the desired set of potential MCP sequences. Methods for synthesizing degenerate oligonucleotides are known in the art (see, e.g., Narang, S.A. (1983) *Tetrahedron* 39:3; Itakura et al. (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al. (1984) *Science* 198:1056; Ike et al. (1983) *Nucleic Acid Res.* 11:477.

In addition, libraries of fragments of the MCP protein coding can be used to 30 generate a variegated population of MCP fragments for screening and subsequent selection of homologues of an MCP protein. In one embodiment, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of an MCP coding sequence with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to 35 form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression

vector. By this method, an expression library can be derived which encodes N-terminal, C-terminal and internal fragments of various sizes of the MCP protein.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. Such techniques are adaptable for rapid screening of the gene libraries generated by the combinatorial mutagenesis of MCP homologues. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a new technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify MCP homologues (Arkin and Yourvan (1992) *PNAS* 89:7811-7815; Delgrave et al. (1993) *Protein Engineering* 6(3):327-331).

In another embodiment, cell based assays can be exploited to analyze a variegated MCP library, using methods well known in the art.

20 *D. Uses and Methods of the Invention*

The nucleic acid molecules, proteins, protein homologues, fusion proteins, primers, vectors, and host cells described herein can be used in one or more of the following methods: identification of *C. glutamicum* and related organisms; mapping of genomes of organisms related to *C. glutamicum*; identification and localization of *C. glutamicum* sequences of interest; evolutionary studies; determination of MCP protein regions required for function; modulation of an MCP protein activity; modulation of the activity of one or more metabolic pathways; and modulation of cellular production of a desired compound, such as a fine chemical.

30 The MCP nucleic acid molecules of the invention have a variety of uses. First, they may be used to identify an organism as being *Corynebacterium glutamicum* or a close relative thereof. Also, they may be used to identify the presence of *C. glutamicum* or a relative thereof in a mixed population of microorganisms. The invention provides the nucleic acid sequences of a number of *C. glutamicum* genes, and probes based thereon; by probing the extracted genomic DNA of a culture of a unique or mixed population of microorganisms under stringent conditions with a probe spanning a region of a *C. glutamicum* gene which is unique to this organism, one can ascertain whether this organism is present. Although *Corynebacterium glutamicum* itself is

nonpathogenic, it is related to pathogenic species, such as *Corynebacterium diphtheriae*. Detection of such organisms is of significant clinical relevance.

To detect the presence of *C. glutamicum* in a sample, techniques well known in the art may be employed. Specifically, the cells in the sample may optionally first be

5 cultured in a suitable liquid or on a suitable solid culture medium to increase the number of cells in the sample. These cells are lysed, and the total DNA content extracted and optionally purified to remove debris and protein material which may interfere with subsequent analysis. The polymerase chain reaction or a similar technique known in the art is performed (for general reference on methodologies commonly used for the

10 amplification of nucleic acid sequences, see Mullis et al., U.S. Patent No. 4,683,195. Mullis et al., U.S. Patent No. 4,965,188, and Innis, M.A., and Gelfand, D. H., (1989) PCR Protocols. A guide to Methods and Applications, Academic Press, p. 3-12, and (1988) Biotechnology 6:1197, and International Patent Application No. WO89/01050) in which primers specific to an MCP nucleic acid molecule of the invention are

15 incubated with the nucleic acid sample such that, if present in the sample, that particular MCP nucleic acid sequence will be amplified. The particular MCP nucleic acid to be amplified is selected based on its uniqueness to the *C. glutamicum* genome, or to the genomes of *C. glutamicum* and only a few closely related bacteria. The presence of the desired amplified product is thus indicative of the presence of *C. glutamicum*, or an

20 organism closely related to *C. glutamicum*.

Further, the nucleic acid and protein molecules of the invention may serve as markers for specific regions of the genome. It is possible, using techniques well known in the art, to ascertain the physical location on the *C. glutamicum* genome of the MCP nucleic acid molecules of the invention, which in turn provides markers on the genome

25 which can be used to aid in the placement of other nucleic acid molecules and genes on the genome map. Also, the nucleic acid molecules of the invention may be sufficiently homologous to the sequences of related bacterial species that these nucleic acid molecules may similarly permit the construction of a genomic map in such bacteria (e.g., *Brevibacterium lactofermentum*).

30 The nucleic acid molecules of the invention have utility not only in the mapping of the genome, but also for functional studies of *C. glutamicum* proteins. For example, to identify the region of the genome to which a particular *C. glutamicum* DNA-binding protein binds, the *C. glutamicum* genome could be digested, and the fragments incubated with the DNA-binding protein. Those which bind the protein may be additionally probed

35 with the nucleic acid molecules of the invention, preferably with readily detectable labels; binding of such a nucleic acid molecule to the genome fragment enables the localization of the fragment to the genome map of *C. glutamicum*, and, when performed

multiple times with different enzymes, facilitates a rapid determination of the nucleic acid sequence to which the protein binds.

The MCP nucleic acid molecules of the invention are also useful for evolutionary and protein structural studies. The metabolic processes in which the molecules of the invention participate are utilized by a wide variety of prokaryotic and eukaryotic cells; by comparing the sequences of the nucleic acid molecules of the present invention to those encoding similar enzymes from other organisms, the evolutionary relatedness of the organisms can be assessed. Similarly, such a comparison permits an assessment of which regions of the sequence are conserved and which are not, which may aid in determining those regions of the protein which are essential for the functioning of the enzyme. This type of determination is of value for protein engineering studies and may give an indication of what the protein can tolerate in terms of mutagenesis without losing function.

The MCP protein molecules of the invention may also be utilized as markers for the classification of an unknown bacterium as *C. glutamicum*, or for the identification of *C. glutamicum* or closely related bacteria in a sample. For example, using techniques well known in the art, cells in a sample may optionally be amplified (e.g., by culturing in an appropriate medium) to increase the sample size, and then may be lysed to release proteins contained therein. This sample may optionally be purified to remove debris and nucleic acid molecules which may interfere with subsequent analysis. Antibodies specific for a selected MCP protein of the invention may be incubated with the protein sample in a typical Western assay format (see, e.g., Ausubel et al., (1988) *Current Protocols in Molecular Biology*, Wiley: New York) in which the antibody will bind to its target protein if this protein is present in the sample. An MCP protein is selected for this type of assay if it is unique or nearly unique to *C. glutamicum* or *C. glutamicum* and bacteria very closely related to *C. glutamicum*. Proteins in the sample are then separated by gel electrophoresis, and transferred to a suitable matrix, such as nitrocellulose. An appropriate secondary antibody having a detectable label (e.g., chemiluminescent or colorimetric) is incubated with this matrix, followed by stringent washing. The presence or absence of the label is indicative of the presence or absence of the target protein in the sample. If the protein is present, then this is indicative of the presence of *C. glutamicum*. A similar process enables the classification of an unknown bacterium as *C. glutamicum*; if a panel of proteins specific to *C. glutamicum* are not detected in protein samples prepared from the unknown bacterium, then that bacterium is not likely to be *C. glutamicum*.

Genetic manipulation of the MCP nucleic acid molecules of the invention may result in the production of MCP proteins having functional differences from the wild-

type MCP proteins. These proteins may be improved in efficiency or activity, may be present in greater numbers in the cell than is usual, or may be decreased in efficiency or activity.

Such changes in activity may directly modulate the yield, production, and/or efficiency of production of one or more fine chemicals from *C. glutamicum*. For example, by modifying the activity of a protein involved in the biosynthesis or degradation of a fine chemical (i.e., through mutagenesis of the corresponding gene), one may directly modulate the ability of the cell to synthesize or to degrade this compound, thereby modulating the yield and/or efficiency of production of the fine chemical. Similarly, by modulating the activity of a protein which regulates a fine chemical metabolic pathway, one may directly influence whether the production of the desired compound is up- or down-regulated, either of which will modulate the yield or efficiency of production of the fine chemical from the cell.

Indirect modulation of fine chemical production may also result by modifying the activity of a protein of the invention (i.e., by mutagenesis of the corresponding gene) such that the overall ability of the cell to grow and divide or to remain viable and productive is increased. The production of fine chemicals from *C. glutamicum* is generally accomplished by the large-scale fermentative culture of these microorganisms, conditions which are frequently suboptimal for growth and cell division. By engineering a protein of the invention (e.g., a stress response protein, a cell wall protein, or proteins involved in the metabolism of compounds necessary for cell growth and division to occur, such as nucleotides and amino acids) such that it is better able to survive, grow, and multiply in such conditions, it may be possible to increase the number and productivity of such engineered *C. glutamicum* cells in large-scale culture, which in turn should result in increased yields and/or efficiency of production of one or more desired fine chemicals. Further, the metabolic pathways of any cell are necessarily interrelated and coregulated. By altering the activity or regulation of any one metabolic pathway in *C. glutamicum* (i.e., by altering the activity of one of the proteins of the invention which participates in such a pathway), it is possible to concomitantly alter the activity or regulation of other metabolic pathways in this microorganism, which may be directly involved in the synthesis or degradation of a fine chemical.

The aforementioned mutagenesis strategies for MCP proteins to result in increased yields of a fine chemical from *C. glutamicum* are not meant to be limiting; variations on these strategies will be readily apparent to one skilled in the art. Using such strategies, and incorporating the mechanisms disclosed herein, the nucleic acid and protein molecules of the invention may be utilized to generate *C. glutamicum* or related strains of bacteria expressing mutated MCP nucleic acid and protein molecules such that

the yield, production, and/or efficiency of production of a desired compound is improved. This desired compound may be any natural product of *C. glutamicum*, which includes the final products of biosynthesis pathways and intermediates of naturally-occurring metabolic pathways, as well as molecules which do not naturally occur in the 5 metabolism of *C. glutamicum*, but which are produced by a *C. glutamicum* strain of the invention.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patent applications, patents, and published patent applications cited throughout this application are hereby incorporated by reference.

5

Exemplification

Example 1: Preparation of total genomic DNA of *Corynebacterium glutamicum* ATCC 13032

10 A culture of *Corynebacterium glutamicum* (ATCC 13032) was grown overnight at 30°C with vigorous shaking in BHI medium (Difco). The cells were harvested by centrifugation. the supernatant was discarded and the cells were resuspended in 5 ml buffer-I (5% of the original volume of the culture — all indicated volumes have been calculated for 100 ml of culture volume). Composition of buffer-I: 140.34 g/l sucrose, 15 2.46 g/l MgSO₄ x 7H₂O, 10 ml/l KH₂PO₄ solution (100 g/l, adjusted to pH 6.7 with KOH), 50 ml/l M12 concentrate (10 g/l (NH₄)₂SO₄, 1 g/l NaCl, 2 g/l MgSO₄ x 7H₂O, 0.2 g/l CaCl₂, 0.5 g/l yeast extract (Difco), 10 ml/l trace-elements-mix (200 mg/l FeSO₄ x H₂O, 10 mg/l ZnSO₄ x 7 H₂O, 3 mg/l MnCl₂ x 4 H₂O, 30 mg/l H₃BO₃, 20 mg/l CoCl₂ x 6 H₂O, 1 mg/l NiCl₂ x 6 H₂O, 3 mg/l Na₂MoO₄ x 2 H₂O, 500 mg/l complexing agent 20 (EDTA or citric acid), 100 ml/l vitamins-mix (0.2 mg/l biotin, 0.2 mg/l folic acid, 20 mg/l p-amino benzoic acid, 20 mg/l riboflavin, 40 mg/l ca-panthothenate, 140 mg/l nicotinic acid, 40 mg/l pyridoxole hydrochloride, 200 mg/l myo-inositol). Lysozyme was added to the suspension to a final concentration of 2.5 mg/ml. After an approximately 4 h incubation at 37°C, the cell wall was degraded and the resulting 25 protoplasts are harvested by centrifugation. The pellet was washed once with 5 ml buffer-I and once with 5 ml TE-buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8). The pellet was resuspended in 4 ml TE-buffer and 0.5 ml SDS solution (10%) and 0.5 ml NaCl solution (5 M) are added. After adding of proteinase K to a final concentration of 200 µg/ml, the suspension is incubated for ca. 18 h at 37°C. The DNA was purified by 30 extraction with phenol, phenol-chloroform-isoamylalcohol and chloroform-isoamylalcohol using standard procedures. Then, the DNA was precipitated by adding 1/50 volume of 3 M sodium acetate and 2 volumes of ethanol, followed by a 30 min incubation at -20°C and a 30 min centrifugation at 12,000 rpm in a high speed centrifuge using a SS34 rotor (Sorvall). The DNA was dissolved in 1 ml TE-buffer containing 20

μg/ml RNaseA and dialysed at 4°C against 1000 ml TE-buffer for at least 3 hours. During this time, the buffer was exchanged 3 times. To aliquots of 0.4 ml of the dialysed DNA solution, 0.4 ml of 2 M LiCl and 0.8 ml of ethanol are added. After a 30 min incubation at -20°C, the DNA was collected by centrifugation (13.000 rpm, Biofuge Fresco, Heraeus, Hanau, Germany). The DNA pellet was dissolved in TE-buffer. DNA prepared by this procedure could be used for all purposes, including southern blotting or construction of genomic libraries.

Example 2: Construction of genomic libraries in *Escherichia coli* of *Corynebacterium glutamicum* ATCC13032.

Starting from DNA prepared as described in Example 1, cosmid and plasmid libraries were constructed according to known and well established methods (see e.g., Sambrook, J. et al. (1989) "Molecular Cloning : A Laboratory Manual", Cold Spring Harbor Laboratory Press, or Ausubel, F.M. et al. (1994) "Current Protocols in Molecular Biology", John Wiley & Sons.)

Any plasmid or cosmid could be used. Of particular use were the plasmids pBR322 (Sutcliffe, J.G. (1979) Proc. Natl. Acad. Sci. USA. 75:3737-3741); pACYC177 (Change & Cohen (1978) J. Bacteriol. 134:1141-1156), plasmids of the pBS series (pBSSK+, pBSSK- and others; Stratagene, LaJolla, USA), or cosmids as SuperCos1 (Stratagene, LaJolla, USA) or Lorist6 (Gibson, T.J., Rosenthal A. and Waterson, R.H. (1987) Gene 53:283-286.

Example 3: DNA Sequencing and Computational Functional Analysis

Genomic libraries as described in Example 2 were used for DNA sequencing according to standard methods, in particular by the chain termination method using ABI377 sequencing machines (see e.g., Fleischman, R.D. et al. (1995) "Whole-genome Random Sequencing and Assembly of *Haemophilus Influenzae* Rd.", Science, 269:496-512). Sequencing primers with the following nucleotide sequences were used: 5'-GGAAACAGTATGACCATG-3' or 5'-GTAAAACGACGGCCAGT-3'.

30 Example 4: *In vivo* Mutagenesis

In vivo mutagenesis of *Corynebacterium glutamicum* can be performed by passage of plasmid (or other vector) DNA through *E. coli* or other microorganisms (e.g. *Bacillus* spp. or yeasts such as *Saccharomyces cerevisiae*) which are impaired in their capabilities to maintain

the integrity of their genetic information. Typical mutator strains have mutations in the genes for the DNA repair system (e.g., mutHLS, mutD, mutT, etc.; for reference, see Rupp, W.D. (1996) DNA repair mechanisms, in: *Escherichia coli* and *Salmonella*, p. 2277-2294. ASM: Washington.) Such strains are well known to those skilled in the art. The use of such strains is 5 illustrated, for example, in Greener, A. and Callahan, M. (1994) Strategies 7: 32-34.

Example 5: DNA Transfer Between *Escherichia coli* and *Corynebacterium glutamicum*

Several *Corynebacterium* and *Brevibacterium* species contain endogenous 10 plasmids (as e.g., pHM1519 or pBL1) which replicate autonomously (for review see, e.g., Martin, J.F. et al. (1987) Biotechnology, 5:137-146). Shuttle vectors for *Escherichia coli* and *Corynebacterium glutamicum* can be readily constructed by using standard vectors for *E. coli* (Sambrook, J. et al. (1989). "Molecular Cloning: A Laboratory Manual". Cold Spring Harbor Laboratory Press or Ausubel, F.M. et al. (1994) "Current Protocols in 15 Molecular Biology". John Wiley & Sons) to which a origin or replication for and a suitable marker from *Corynebacterium glutamicum* is added. Such origins of replication are preferably taken from endogenous plasmids isolated from *Corynebacterium* and *Brevibacterium* species. Of particular use as transformation markers for these species are genes for kanamycin resistance (such as those derived from the Tn5 or Tn903 20 transposons) or chloramphenicol (Winnacker, E.L. (1987) "From Genes to Clones — Introduction to Gene Technology, VCH, Weinheim). There are numerous examples in the literature of the construction of a wide variety of shuttle vectors which replicate in both *E. coli* and *C. glutamicum*, and which can be used for several purposes, including gene over-expression (for reference, see e.g., Yoshihama, M. et al. (1985) J. Bacteriol. 162:591-597, 25 Martin J.F. et al. (1987) Biotechnology, 5:137-146 and Eikmanns, B.J. et al. (1991) Gene. 102:93-98).

Using standard methods, it is possible to clone a gene of interest into one of the shuttle vectors described above and to introduce such a hybrid vectors into strains of *Corynebacterium glutamicum*. Transformation of *C. glutamicum* can be achieved by 30 protoplast transformation (Kastsumata, R. et al. (1984) J. Bacteriol. 159:306-311), electroporation (Liebl, E. et al. (1989) FEMS Microbiol. Letters. 53:399-303) and in cases where special vectors are used, also by conjugation (as described e.g. in Schäfer, A et al. (1990) J. Bacteriol. 172:1663-1666). It is also possible to transfer the shuttle vectors for

C. glutamicum to *E. coli* by preparing plasmid DNA from *C. glutamicum* (using standard methods well-known in the art) and transforming it into *E. coli*. This transformation step can be performed using standard methods, but it is advantageous to use an *Mcr*-deficient *E. coli* strain, such as NM522 (Gough & Murray (1983) *J. Mol. Biol.* 166:1-19).

5

Example 6: Assessment of the Expression of the Mutant Protein

Observations of the activity of a mutated protein in a transformed host cell rely on the fact that the mutant protein is expressed in a similar fashion and in a similar quantity to that of the wild-type protein. A useful method to ascertain the level of transcription of 10 the mutant gene (an indicator of the amount of mRNA available for translation to the gene product) is to perform a Northern blot (for reference see, for example, Ausubel et al. (1988) *Current Protocols in Molecular Biology*, Wiley: New York), in which a primer designed to bind to the gene of interest is labeled with a detectable tag (usually radioactive or chemiluminescent), such that when the total RNA of a culture of the organism is 15 extracted, run on gel, transferred to a stable matrix and incubated with this probe, the binding and quantity of binding of the probe indicates the presence and also the quantity of mRNA for this gene. This information is evidence of the degree of transcription of the mutant gene. Total cellular RNA can be prepared from *Corynebacterium glutamicum* by several methods, all well-known in the art, such as that described in Bormann, E.R. et al. 20 (1992) *Mol. Microbiol.* 6: 317-326.

To assess the presence or relative quantity of protein translated from this mRNA, standard techniques, such as a Western blot, may be employed (see, for example, Ausubel et al. (1988) *Current Protocols in Molecular Biology*, Wiley: New York). In this process, total cellular proteins are extracted, separated by gel electrophoresis, transferred to a 25 matrix such as nitrocellulose, and incubated with a probe, such as an antibody, which specifically binds to the desired protein. This probe is generally tagged with a chemiluminescent or colorimetric label which may be readily detected. The presence and quantity of label observed indicates the presence and quantity of the desired mutant protein present in the cell.

30

Example 7: Growth of Genetically Modified *Corynebacterium glutamicum* — Media and Culture Conditions

Genetically modified *Corynebacteria* are cultured in synthetic or natural growth media. A number of different growth media for *Corynebacteria* are both well-known and readily available (Lieb *et al.* (1989) *Appl. Microbiol. Biotechnol.*, 32:205-210; von der Osten *et al.* (1998) *Biotechnology Letters*, 11:11-16; Patent DE 4,120.867; Liebl (1992) "The Genus *Corynebacterium*, in: *The Prokaryotes*, Volume II, Balows, A. *et al.*, eds. Springer-Verlag). These media consist of one or more carbon sources, nitrogen sources, inorganic salts, vitamins and trace elements. Preferred carbon sources are sugars, such as mono-, di-, or polysaccharides. For example, glucose, fructose, mannose, galactose, ribose, sorbose, ribulose, lactose, maltose, sucrose, raffinose, starch or cellulose serve as very good carbon sources. It is also possible to supply sugar to the media via complex compounds such as molasses or other by-products from sugar refinement. It can also be advantageous to supply mixtures of different carbon sources. Other possible carbon sources are alcohols and organic acids, such as methanol, ethanol, acetic acid or lactic acid. Nitrogen sources are usually organic or inorganic nitrogen compounds, or materials which contain these compounds. Exemplary nitrogen sources include ammonia gas or ammonia salts, such as NH₄Cl or (NH₄)₂SO₄, NH₄OH, nitrates, urea, amino acids or complex nitrogen sources like corn steep liquor, soy bean flour, soy bean protein, yeast extract, meat extract and others.

Inorganic salt compounds which may be included in the media include the chloride-, phosphorous- or sulfate- salts of calcium, magnesium, sodium, cobalt, molybdenum, potassium, manganese, zinc, copper and iron. Chelating compounds can be added to the medium to keep the metal ions in solution. Particularly useful chelating compounds include dihydroxyphenols, like catechol or protocatechuate, or organic acids, such as citric acid. It is typical for the media to also contain other growth factors, such as vitamins or growth promoters, examples of which include biotin, riboflavin, thiamin, folic acid, nicotinic acid, pantothenate and pyridoxin. Growth factors and salts frequently originate from complex media components such as yeast extract, molasses, corn steep liquor and others. The exact composition of the media compounds depends strongly on the immediate experiment and is individually decided for each specific case. Information about media optimization is available in the textbook "Applied Microbiol. Physiology, A Practical Approach (eds. P.M. Rhodes, P.F. Stanbury, IRL Press (1997) pp. 53-73, ISBN 0

19 963577 3). It is also possible to select growth media from commercial suppliers, like standard 1 (Merck) or BHI (grain heart infusion, DIFC) or others.

All medium components are sterilized, either by heat (20 minutes at 1.5 bar and 121°C) or by sterile filtration. The components can either be sterilized together or, if 5 necessary, separately. All media components can be present at the beginning of growth, or they can optionally be added continuously or batchwise.

Culture conditions are defined separately for each experiment. The temperature should be in a range between 15°C and 45°C. The temperature can be kept constant or can be altered during the experiment. The pH of the medium should be in the range of 5 to 10 8.5, preferably around 7.0, and can be maintained by the addition of buffers to the media. An exemplary buffer for this purpose is a potassium phosphate buffer. Synthetic buffers such as MOPS, HEPES, ACES and others can alternatively or simultaneously be used. It is also possible to maintain a constant culture pH through the addition of NaOH or NH₄OH during growth. If complex medium components such as yeast extract are utilized, 15 the necessity for additional buffers may be reduced, due to the fact that many complex compounds have high buffer capacities. If a fermentor is utilized for culturing the micro-organisms, the pH can also be controlled using gaseous ammonia.

The incubation time is usually in a range from several hours to several days. This time is selected in order to permit the maximal amount of product to accumulate in the 20 broth. The disclosed growth experiments can be carried out in a variety of vessels, such as microtiter plates, glass tubes, glass flasks or glass or metal fermentors of different sizes. For screening a large number of clones, the microorganisms should be cultured in microtiter plates, glass tubes or shake flasks, either with or without baffles. Preferably 25 100 ml shake flasks are used, filled with 10% (by volume) of the required growth medium. The flasks should be shaken on a rotary shaker (amplitude 25 mm) using a speed-range of 100 – 300 rpm. Evaporation losses can be diminished by the maintenance of a humid atmosphere; alternatively, a mathematical correction for evaporation losses should be performed.

If genetically modified clones are tested, an unmodified control clone or a control 30 clone containing the basic plasmid without any insert should also be tested. The medium is inoculated to an OD₆₀₀ of 0.5 – 1.5 using cells grown on agar plates, such as CM plates (10 g/l glucose, 2.5 g/l NaCl, 2 g/l urea, 10 g/l polypeptone, 5 g/l yeast extract, 5 g/l meat extract, 22 g/l NaCl, 2 g/l urea, 10 g/l polypeptone, 5 g/l yeast extract, 5 g/l meat extract,

22 g/l agar, pH 6.8 with 2M NaOH) that had been incubated at 30°C. Inoculation of the media is accomplished by either introduction of a saline suspension of *C. glutamicum* cells from CM plates or addition of a liquid preculture of this bacterium.

5 **Example 8 – *In vitro* Analysis of the Function of Mutant Proteins**

The determination of activities and kinetic parameters of enzymes is well established in the art. Experiments to determine the activity of any given altered enzyme must be tailored to the specific activity of the wild-type enzyme, which is well within the ability of one skilled in the art. Overviews about enzymes in general, as well as specific details concerning structure, kinetics, principles, methods, applications and examples for the determination of many enzyme activities may be found, for example, in the following references: Dixon, M. and Webb, E.C., (1979) *Enzymes*. Longmans: London; Fersht, (1985) *Enzyme Structure and Mechanism*. Freeman: New York; Walsh, (1979) *Enzymatic Reaction Mechanisms*. Freeman: San Francisco; Price, N.C., Stevens, L. (1982) *Fundamentals of Enzymology*. Oxford Univ. Press: Oxford; Boyer, P.D., ed. (1983) *The Enzymes*, 3rd ed. Academic Press: New York; Bisswanger, H., (1994) *Enzymkinetik*, 2nd ed. VCH: Weinheim (ISBN 3527300325); Bergmeyer, H.U., Bergmeyer, J., Graßl, M., eds. (1983-1986) *Methods of Enzymatic Analysis*, 3rd ed., vol. I-XII, Verlag Chemie: Weinheim; and Ullmann's *Encyclopedia of Industrial Chemistry* (1987) vol. A9, "Enzymes". VCH: Weinheim, p. 352-363.

The activity of proteins which bind to DNA can be measured by several well-established methods, such as DNA band-shift assays (also called gel retardation assays). The effect of such proteins on the expression of other molecules can be measured using reporter gene assays (such as that described in Kolmar, H. et al. (1995) *EMBO J.* 14: 3895-3904 and references cited therein). Reporter gene test systems are well known and established for applications in both pro- and eukaryotic cells, using enzymes such as beta-galactosidase, green fluorescent protein, and several others.

The determination of activity of membrane-transport proteins can be performed according to techniques such as those described in Gennis, R.B. (1989) "Pores, Channels and Transporters", in *Biomembranes, Molecular Structure and Function*, Springer: Heidelberg, p. 85-137; 199-234; and 270-322.

Example 9: Analysis of Impact of Mutant Protein on the Production of the Desired Product

35 The effect of the genetic modification in *C. glutamicum* on production of a desired compound (such as an amino acid) can be assessed by growing the modified microorganism under suitable conditions (such as those described above) and analyzing

the medium and/or the cellular component for increased production of the desired product (i.e., an amino acid). Such analysis techniques are well known to one skilled in the art, and include spectroscopy, thin layer chromatography, staining methods of various kinds, enzymatic and microbiological methods, and analytical chromatography such as high performance liquid chromatography (see, for example, Ullman, Encyclopedia of Industrial Chemistry, vol. A2, p. 89-90 and p. 443-613, VCH: Weinheim (1985); Fallon, A. et al., (1987) "Applications of HPLC in Biochemistry" in: Laboratory Techniques in Biochemistry and Molecular Biology, vol. 17; Rehm et al. (1993) Biotechnology, vol. 3, Chapter III: "Product recovery and purification", page 469-714, VCH: Weinheim; Belter, P.A. et al. (1988) Bioseparations: downstream processing for biotechnology, John Wiley and Sons; Kennedy, J.F. and Cabral, J.M.S. (1992) Recovery processes for biological materials, John Wiley and Sons; Shaeiwitz, J.A. and Henry, J.D. (1988) Biochemical separations, in: Ullmann's Encyclopedia of Industrial Chemistry, vol. B3, Chapter 11, page 1-27. VCH: Weinheim; and Dechow, F.J. (1989) Separation and purification techniques in biotechnology, Noyes Publications.)

In addition to the measurement of the final product of fermentation, it is also possible to analyze other components of the metabolic pathways utilized for the production of the desired compound, such as intermediates and side-products, to determine the overall efficiency of production of the compound. Analysis methods include measurements of nutrient levels in the medium (e.g., sugars, hydrocarbons, nitrogen sources, phosphate, and other ions), measurements of biomass composition and growth, analysis of the production of common metabolites of biosynthetic pathways, and measurement of gasses produced during fermentation. Standard methods for these measurements are outlined in Applied Microbial Physiology, A Practical Approach, P.M. Rhodes and P.F. Stanbury, eds., IRL Press, p. 103-129: 131-163: and 165-192 (ISBN: 0199635773) and references cited therein.

Example 10: Purification of the Desired Product from *C. glutamicum* Culture

Recovery of the desired product from the *C. glutamicum* cells or supernatant of the above-described culture can be performed by various methods well known in the art. If the desired product is not secreted from the cells, the cells can be harvested from the culture by low-speed centrifugation, the cells can be lysed by standard techniques, such as mechanical force or sonication. The cellular debris is removed by centrifugation, and the supernatant fraction containing the soluble proteins is retained for further purification of the desired compound. If the product is secreted from the *C. glutamicum*

cells, then the cells are removed from the culture by low-speed centrifugation, and the supernate fraction is retained for further purification.

The supernatant fraction from either purification method is subjected to chromatography with a suitable resin, in which the desired molecule is either retained on

5 a chromatography resin while many of the impurities in the sample are not, or where the impurities are retained by the resin while the sample is not. Such chromatography steps may be repeated as necessary, using the same or different chromatography resins. One skilled in the art would be well-versed in the selection of appropriate chromatography resins and in their most efficacious application for a particular molecule to be purified.

10 The purified product may be concentrated by filtration or ultrafiltration, and stored at a temperature at which the stability of the product is maximized.

There are a wide array of purification methods known to the art and the preceding method of purification is not meant to be limiting. Such purification techniques are described, for example, in Bailey, J.E. & Ollis, D.F. *Biochemical Engineering Fundamentals*, McGraw-Hill: New York (1986).

15 The identity and purity of the isolated compounds may be assessed by techniques standard in the art. These include high-performance liquid chromatography (HPLC), spectroscopic methods, staining methods, thin layer chromatography, NIRS, enzymatic assay, or microbiologically. Such analysis methods are reviewed in: Patek et al. (1994) *Appl. Environ. Microbiol.* 60: 133-140; Malakhova et al. (1996) *Biotehnologiya* 11: 27-32; and Schmidt et al. (1998) *Bioprocess Engineer.* 19: 67-70. Ulmann's Encyclopedia of Industrial Chemistry, (1996) vol. A27, VCH: Weinheim, p. 89-90, p. 521-540, p. 540-547, p. 559-566, 575-581 and p. 581-587; Michal, G. (1999) *Biochemical Pathways: An Atlas of Biochemistry and Molecular Biology*, John Wiley and Sons; Fallon, A. et al. (1987) Applications of HPLC in Biochemistry in: *Laboratory Techniques in Biochemistry and Molecular Biology*, vol. 17.

Equivalents

Those skilled in the art will recognize, or will be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

TABLE 1: GENES IN THE APPLICATION

Identification Code	Config.	NT Start	NT Stop
RXA02223	GR00652	425	6
RXA00911	GR00248	1259	1765
RXA02032	GR00618	4160	4729
RXA01707	GR00481	802	1629
RXA00211	GR00041	3709	2720
RXA02427	GR00707	1447	3061
RXA00499	GR00087	830	1144
RXA01186	GR00338	3742	2645
RXA00150	GR00023	4085	4058
RXA00318	GR00055	426	635
RXA00318	GR00059	1	783
RXA00555	GR00145	2555	1665
RXA00557	GR00169	10882	9980
RXA00910	GR00253	3841	3089
RXA01198	GR00343	3422	3724
RXA01588	GR00443	497	33
RXA01693	GR00474	1553	2974
RXA02425	GR00707	1	630
RXA02573	GR00739	594	151
RXA02865	GR00753	6497	6018
RXA00989	GR00242	15341	15928
RXA02808	GR00787	48	570
RXA01658	GR00460	1548	2444
RXA02721	GR00759	1173	616
RXA00462	GR00116	3023	1644
RXA01286	GR00387	14457	13423
RXA01380	GR00403	2	2017
RXA02529	GR00725	7943	8071
RXA00227	GR00003	5142	5507
RXA00117	GR00019	791	201
RXA02447	GR00037	7097	6171
RXA01815	GR00515	3294	4085
RXA02138	GR00639	4409	4750
RXA02107	GR00632	1536	1877
RXA02180	GR00841	16813	15356
RXA01986	GR00567	47	703
RXA0411	GR00092	1685	1011
RXA01982	GR00573	3001	1844

Identification Code	Config.	Start	NT	Stop
RXA02367	GR006807	2182	1554	
RXA02884	GR10020	1695	2158	
RXA02713	GR00762	6407	6027	
RXA02840	GR00825	488	139	
RXA01998	GR00585	88	624	
RXA01195	GR00443	1413	1839	
RXA00305	GR00051	1257	826	
RXA02383	GR00692	608	6	
RXA02715	GR00763	777	73	
RXA00239	GR00038	5118	4534	
RXA01091	GR00305	546	76	
RXA02690	GR00754	14502	13405	
RXA08667	GR00175	593	1177	
RXA00356	GR00070	2853	104	
RXA01628	GR00185	1284	877	
RXA01719	GR00188	5283	6911	
RXA01845	GR00456	10574	9989	
RXA02070	GR00627	1733	2830	
RXA00349	GR00966	3	1061	
RXA02124	GR00868	1548	2633	
RXA00848	GR00948	113	511	
RXA00153	GR00023	7656	7231	
RXA00417	GR00093	404	3100	
RXA02443	GR00179	6818	7771	
RXA00325	GR00057	8594	9238	
RXA00874	GR00241	758	1848	
RXA02403	GR00700	896	1860	
RXA01271	GR00387	23467	21656	
RXA01268	GR00367	19165	18526	
RXA01846	GR00458	11513	10695	
RXA01671	GR00486	854	1468	
RXA00805	GR00215	2057	2938	
RXA00008	GR00002	605	115	
RXA01359	GR00193	6887	8038	
RXA00881	GR00225	6	431	
RXA01076	GR00300	4374	3355	
RXA02244	GR00654	12058	13580	
RXA01696	GR00475	799	203	
RXA02545	GR010728	18149	18192	
RXA02688	GR00754	12258	12924	
RXA02689	GR00754	13405	13084	
RXA02588	GR00741	13037	12254	
RXA01367	GR00397	1518	1919	
RXA01577	GR00438	8811	9165	
RXA01585	GR00441	1226	800	
RXA01492	GR00423	6133	6330	
RXA01592	GR00447	3	1295	

Identification Code	Config.	Start	NT	NT Stop
RXA01597	GR00447	6220	7401	
RXA01176	GR00335	1980	1477	
RXA01748	GR00495	3681	4460	
RXA02137	GR00839	4166	3369	
RXA02141	GR00839	8457	8864	
RXA02076	GR00628	6902	7455	
RXA00473	GR00119	5769	6561	
RXA00211	GR00036	420	4	
RXA00224	GR00036	998	459	
RXA00161	GR00024	4893	5334	
RXA00183	GR00028	7344	8195	
RXA00279	GR00043	4001	2816	
RXA00474	GR00119	6575	8152	
RXA02314	GR00685	8179	5819	
RXA00560	GR00149	256	492	
RXA00587	GR00156	13008	13490	
RXA02575	GR00719	1907	3064	
RXA02824	GR00805	531	4	
RXA02849	GR00849	2	283	
RXA01159	GR00328	3089	2775	
RXA01023	GR00292	1817	867	
RXA01944	GR00558	2	385	
RXA01635	GR00454	5575	6315	
RXA01638	GR00454	6326	8898	
RXA01945	GR00558	392	1633	
RXA01968	GR00587	3295	2138	
RXA02452	GR00710	5271	5092	
RXA02183	GR00541	18663	19187	
RXA00614	GR00162	1680	2594	
RXA01322	GR00385	443	6	
RXA01142	GR00389	11296	12807	
RXA00054	GR00009	8557	11469	
RXA00096	GR00114	4148	5048	
RXA00097	GR00114	5222	6382	
RXA00118	GR00019	918	1172	
RXA00122	GR00019	4220	5842	
RXA00134	GR00021	1648	1079	
RXA00159	GR00024	3868	2687	
RXA00185	GR00028	9418	12045	
RXA00220	GR00032	20866	20163	
RXA00248	GR00037	7843	7121	
RXA00285	GR00046	3	515	
RXA00321	GR00057	2411	597	
RXA00322	GR00057	3659	2555	
RXA00339	GR00059	817	1533	
RXA00398	GR00086	6653	6163	
RXA00422	GR00097	428	6	

Identification	Code	NT	Stop
	Config.	Start	
RXA00428	GR00088	2657	2025
RXA0491	GR00122	1057	638
RXA00505	GR00126	1	252
RXA00540	GR00139	2027	2289
RXA01052	GR00145	2	718
RXA00555	GR00145	742	1082
RXA00573	GR00158	117	767
RXA00574	GR00158	767	1645
RXA00576	GR00156	4087	3365
RXA00586	GR00156	12818	11937
RXA00649	GR00161	1193	2056
RXA00688	GR00162	1652	1200
RXA00691	GR00181	2152	1223
RXA00713	GR00188	71	1033
RXA00716	GR00188	3002	3514
RXA00722	GR00189	1015	512
RXA00738	GR00201	78	365
RXA00765	GR00204	3283	3969
RXA00781	GR00204	5780	5993
RXA00768	GR00204	5958	6389
RXA00781	GR00206	2682	2395
RXA00846	GR00230	391	5
RXA00859	GR00234	4	636
RXA00869	GR00239	1	792
RXA00887	GR00242	1354	14266
RXA00940	GR00257	129	524
RXA00949	GR00259	5400	8047
RXA00946	GR00280	60	401
RXA00947	GR00280	875	411
RXA01011	GR00288	2089	857
RXA01017	GR00290	2175	1587
RXA01021	GR00291	1759	2280
RXA01074	GR00300	2811	2107
RXA01078	GR00300	6043	6876
RXA01088	GR00304	3083	1902
RXA01129	GR00314	1461	3326
RXA01196	GR00343	1889	2578
RXA01197	GR00343	3333	2881
RXA01207	GR00347	126	773
RXA01237	GR00358	2751	2311
RXA01246	GR00380	1824	2462
RXA01249	GR00383	303	4
RXA01251	GR00385	228	536
RXA01282	GR00389	5444	4665
RXA01294	GR00373	3537	2872
RXA01348	GR00392	261	752
RXA01357	GR00393	4357	4659

<u>Identification</u>	<u>Code</u>	<u>Contig.</u>	<u>Start</u>	<u>NT</u>	<u>Stop</u>
RXA01382	GR00395	J	1397	1397	
RXA01384	GR00396	1869	4	4	
RXA01388	GR00397	1389	960	960	
RXA01370	GR00398	1875	2225	2225	
RXA01372	GR00399	1	591	591	
RXA01379	GR00402	928	6	6	
RXA01396	GR00408	8475	6218	6218	
RXA01397	GR00408	6894	6475	6475	
RXA01408	GR00410	5288	4481	4481	
RXA01429	GR00417	5651	6268	6268	
RXA01439	GR00418	5949	6494	6494	
RXA01463	GR00421	2493	1330	1330	
RXA01489	GR00423	2179	1349	1349	
RXA01497	GR00424	262	1179	1179	
RXA01501	GR00424	81140	7843	7843	
RXA01505	GR00424	11318	11815	11815	
RXA01523	GR00424	27851	28901	28901	
RXA01588	GR00447	3328	4285	4285	
RXA01600	GR00447	10480	11128	11128	
RXA01622	GR00452	1908	2510	2510	
RXA01682	GR00462	1890	2432	2432	
RXA01709	GR00483	745	416	416	
RXA01715	GR00485	1267	1962	1962	
RXA01738	GR00493	3971	4884	4884	
RXA01803	GR00509	5671	4712	4712	
RXA01804	GR00509	6117	5797	5797	
RXA01805	GR00509	6515	6186	6186	
RXA01844	GR00522	1950	1771	1771	
RXA01871	GR00534	2797	3759	3759	
RXA01875	GR00536	516	1313	1313	
RXA01877	GR00537	135	1199	1199	
RXA01879	GR00537	2117	2704	2704	
RXA01880	GR00537	2641	3049	3049	
RXA01896	GR00544	2	580	580	
RXA01916	GR00549	1034	2044	2044	
RXA01931	GR00555	4913	5588	5588	
RXA01942	GR00557	1526	2927	2927	
RXA01992	GR00581	709	260	260	
RXA02023	GR00613	3234	4001	4001	
RXA02057	GR00625	2972	3502	3502	
RXA02071	GR00628	458	6	6	
RXA02104	GR00631	5327	4908	4908	
RXA02108	GR00632	2077	2511	2511	
RXA02117	GR00636	1056	1529	1529	
RXA02123	GR00636	6558	7928	7928	
RXA02124	GR00636	7956	9911	9911	
RXA02166	GR00640	13048	13224	13224	
RXA02177	GR00641	12683	11615	11615	
RXA02187	GR00641	21249	23447	23447	
RXA02211	GR00648	2537	2989	2989	

Table I, Page 5

Identification Code	Config.	Start	NT	Stop	NT	Start	NT	Stop	NT	Start	NT	Stop
RXA02216	GR00851	2	307									
RXA02217	GR00851	960	306									
RXA02218	GR00851	1299	1565									
RXA02219	GR00851	1578	2983									
RXA02255	GR00654	22507	23442									
RXA02298	GR00682	10310	8652									
RXA02308	GR00884	939	511									
RXA02337	GR00672	2893	3816									
RXA02347	GR00677	509	189									
RXA02349	GR00598	2	733									
RXA02396	GR00698	1309	1031									
RXA02407	GR00701	1580	1885									
RXA02409	GR00702	1248	815									
RXA02430	GR00707	7498	7680									
RXA02459	GR00712	4341	5075									
RXA02472	GR00715	545	5175									
RXA02484	GR00718	2317	1817									
RXA02486	GR00718	3441	4016									
RXA02498	GR00720	10025	9219									
RXA02514	GR00723	1	837									
RXA02518	GR00723	3484	3814									
RXA02521	GR00724	2924	4166									
RXA02525	GR00725	3113	3490									
RXA02540	GR00726	12430	12001									
RXA02601	GR00742	5268	7246									
RXA02617	GR00745	1404	1910									
RXA02639	GR00749	511	1344									
RXA02672	GR00753	12303	13400									
RXA02714	GR00758	14754	14328									
RXA02720	GR00759	631	5									
RXA02761	GR00764	6393	5920									
RXA02766	GR00770	986	594									
RXA02789	GR00777	5237	5782									
RXA02796	GR00778	1648	1100									
RXA02874	GR00115	1348	889									
RXA02901	GR00400	9518	10195									
RXA01504	GR00424	10710	11318									
RXA01508	GR00424	11815	12225									
RXA01647	GR00456	12422	11535									
RXA01798	GR00508	2	484									
RXA02112	GR00638	737	1375									
RXA02254	GR00654	21769	22449									
RXA02482	GR00710	914	105									
RXA02789	GR00780	182	454									
RXA00052	GR00008	7957	7247									

<u>Identification</u>	<u>Code</u>	<u>Confg.</u>	<u>Start</u>	<u>NT</u>	<u>NT</u>	<u>Stop</u>
RXA00180		GR00028	2334	1795		
RXA00163		GR000204	1384	2168		
RXA00926		GR00253	4668	104		
RXA01273		GR00367	26475	25042		
RXA02798		GR00776	2842	4286		
RXA02847		GR00847	598	5		
RXA02898		GR10040	1631	6		
RXA02899		GR10040	2125	1846		
RXA00025		GR00003	2211	3647		
RXA00093		GR00014	204	2426		
RXA00101		GR00014	10514	10107		
RXA00108		GR00015	546	4		
RXA00197		GR00030	1731	2741		
RXA00297		GR00048	2861	3772		
RXA00301		GR00049	1970	2506		
RXA00326		GR00057	19481	19931		
RXA00444		GR00063	6	584		
RXA00416		GR00093	1	327		
RXA00418		GR00094	1	1065		
RXA00418		GR00099	3473	3083		
RXA00430		GR00108	518	817		
RXA00447		GR00113	2	619		
RXA00455		GR00119	25230	23108		
RXA00485		GR00121	2678	1774		
RXA00490		GR00126	489	1829		
RXA00506		GR00131	1	482		
RXA00515		GR00132	599	796		
RXA00520		GR00159	4907	4155		
RXA00602		GR00161	3640	2165		
RXA00611		GR00176	797	6		
RXA00688		GR00177	755	6		
RXA00674		GR00185	2613	142		
RXA00731		GR00224	266	988		
RXA00830		GR00228	1	692		
RXA00835		GR00298	2184	3254		
RXA01068		GR00289	2822	2438		
RXA01071		GR00335	2121	4108		
RXA01102		GR00306	10018	8774		
RXA01119		GR00310	1068	139		
RXA01158		GR00328	2580	1639		
RXA01177		GR00335	2121	4108		
RXA01229		GR00355	2806	3498		
RXA01331		GR00367	1606	1031		
RXA01507		GR00424	12239	12061		
RXA01623		GR00452	2514	3224		
RXA01624		GR00452	3220	3564		
RXA01669		GR00465	1002	271		
RXA01673		GR00467	1807	773		
RXA01685		GR00470	1488	910		

Identification	Code	Contig.	NT	Start	NT	Stop
RXA01149	GR00495	48133	6249			
RXA01306	GR00509	6595	7074			
RXA02080	GR00628	11017	10211			
RXA02172	GR00841	6919	6581			
RXA02295	GR00862	6842	6063			
RXA02297	GR00862	7502	8838			
RXA02390	GR00865	1500	832			
RXA02408	GR00702	832	5			
RXA02486	GR00719	1	369			
RXA02489	GR00719	173	996			
RXA02495	GR00720	9002	6435			
RXA02524	GR00725	2405	3094			
RXA02544	GR00728	18715	18142			
RXA02584	GR00741	6925	8575			
RXA02885	GR00741	9917	8937			
RXA02938	GR00742	2576	3186			
RXA02860	GR00742	5027	3830			
RXA02802	GR00742	7239	7742			
RXA02804	GR00742	6800	10875			
RXA02653	GR00755	1650	4			
RXA02700	GR00757	3507	4742			
RXA02701	GR00757	4838	6145			
RXA00854	GR00169	7213	8478			
RXA01425	GR00417	1701	2585			
RXA02549	GR00728	1331	6			
RXA02579	GR00740	4385	3878			
RXA02580	GR00740	4982	4239			
RXA00806	GR00216	277	5			
RXA00808	GR00217	1029	352			
RXA01318	GR00382	3616	2315			
RXA01677	GR00467	5043	4300			
RXA01648	GR00461	5	1489			
RXA02697	GR00757	1	699			
RXA02719	GR00758	19598	20245			
RXA00003	GR00001	2279	3019			
RXA00015	GR00002	5999	6307			
RXA00018	GR00002	12979	14277			
RXA00020	GR00002	17142	16363			
RXA00021	GR00002	18766	20538			
RXA00022	GR00002	20583	21297			
RXA00028	GR00003	8058	6112			
RXA00031	GR00003	10383	9982			
RXA00038	GR00004	7204	8619			
RXA00037	GR00004	9557	8685			
RXA00039	GR00005	2099	1431			
RXA00040	GR00006	2499	2095			

Identification Code	Config.	NT	Stop
	Config.	NT	Start
RXA00047	GR00008	514	95
RXA00049	GR00008	2270	2956
RXA00058	CR00009	1463	714
RXA00058	CR00009	7394	6831
RXA00059	CR00009	8301	8020
RXA00063	CR00010	1658	1374
RXA00065	CR00010	4140	4412
RXA00067	CR00011	708	223
RXA00088	CR00011	1305	724
RXA00077	CR00012	4228	5589
RXA00079	CR00012	8593	6920
RXA00080	CR00012	7342	6923
RXA00082	CR00012	9019	8456
RXA00083	CR00013	771	1070
RXA00086	CR00013	2739	3092
RXA00087	CR00013	3983	3456
RXA00094	CR00014	3163	3435
RXA00110	CR00016	364	912
RXA00114	CR00017	3420	3908
RXA00119	CR00019	1704	2462
RXA00120	CR00019	2798	3451
RXA00121	CR00019	3473	4183
RXA00127	CR00020	2871	2416
RXA00128	CR00020	4709	3006
RXA00140	CR00022	3841	3656
RXA00141	CR00022	4307	3846
RXA00142	CR00022	4776	4300
RXA00151	CR00023	4956	5552
RXA00154	CR00023	8568	7728
RXA00155	CR00023	8615	9397
RXA00162	CR00024	5438	5191
RXA00167	CR00025	4324	4584
RXA00189	CR00026	5222	3150
RXA00170	CR00026	9914	8081
RXA00171	CR00028	10316	10086
RXA00173	CR00027	1718	1384
RXA00174	CR00027	2079	1795
RXA00175	CR00027	2132	2103
RXA00176	CR00027	1415	3317
RXA00179	CR00028	1714	1258
RXA00194	CR00030	290	6
RXA00199	CR00031	2172	754
RXA00200	CR00031	2837	2535
RXA00207	CR00032	8430	6147
RXA00211	CR00032	10120	10782
RXA00218	CR00032	16104	19243
RXA00222	CR00032	21073	22218
RXA00230	CR00034	746	27

Identification	Code	NT	Start	Stop
		Config.	Start	Stop
RXA00232		GR00035	527	18
RXA00236		GR00036	3300	2575
RXA00237		GR00036	3668	4045
RXA00238		GR00038	4168	4554
RXA00240		GR00036	6342	5133
RXA00242		GR00036	7031	8233
RXA00244		GR00037	1565	930
RXA00245		GR00037	3049	1565
RXA00250		GR00038	6	221
RXA00252		GR00039	485	727
RXA00255		GR00039	2	604
RXA00258		GR00039	988	1738
RXA00257		GR00039	1760	2215
RXA00258		GR00039	3219	3890
RXA00260		GR00039	9214	10409
RXA00261		GR00039	11693	11265
RXA00264		GR00040	2459	2836
RXA00267		GR00040	4091	3822
RXA00276		GR00041	4420	4791
RXA00272		GR00042	185	1297
RXA00273		GR00042	1556	4165
RXA00274		GR00042	4696	4238
RXA00275		GR00042	5016	4875
RXA00277		GR00042	793	5
RXA00282		GR00044	793	
RXA00283		GR00045	142	1269
RXA00286		GR00048	578	1142
RXA00276		GR00047	2781	3189
RXA00294		GR00047	9378	9857
RXA00126		GR00049	2595	3416
RXA00102		GR00050	459	4
RXA00134		GR00050	1081	887
RXA00103		GR00052	358	537
RXA00308		GR00052	9378	17097
RXA00320		GR00057	16782	
RXA00312		GR00057	16782	
RXA00314		GR00057	16782	
RXA00337		GR00058	530	6
RXA00342		GR00061	73	501
RXA00357		GR00061	4013	5464
RXA00358		GR00070	4069	5199
RXA00362		GR00073	2	981
RXA00373		GR00079	342	4
RXA00375		GR00080	549	49
RXA00380		GR00082	836	216
RXA00384		GR00083	395	6
RXA00387		GR00084	1403	591
RXA00390		GR00086	1437	1841
RXA00392		GR00086	3890	3027

Identification	Code	NT	Stop
	Config.	NT	Start
RXA00194	GR00086	5322	4990
RXA00195	GR00086	5417	5716
RXA00197	GR00086	7206	6667
RXA00198	GR00087	1	681
RXA00198	GR00091	642	1088
RXA00198	GR00091	1088	2500
RXA00199	GR00091	909	457
RXA00199	GR00097	1379	909
RXA00202	GR00097	1433	1857
RXA00205	GR00097	4209	3388
RXA00205	GR00116	1282	464
RXA00205	GR00118	1446	1970
RXA00205	GR00118	816	125
RXA00205	GR00119	1647	412
RXA00205	GR00119	5449	4589
RXA00205	GR00119	8922	8163
RXA00205	GR00119	8961	9821
RXA00205	GR00119	17636	18220
RXA00205	GR00120	1	702
RXA00205	GR00123	1	126
RXA00205	GR00123	1776	2177
RXA00205	GR00125	5007	5252
RXA00207	GR00127	1098	244
RXA00209	GR00128	316	140
RXA00210	GR00128	384	914
RXA00210	GR00132	4	516
RXA00210	GR00134	111	575
RXA00222	GR00136	3123	1380
RXA00227	GR00136	3562	4650
RXA00228	GR00136	5274	4712
RXA00229	GR00136	6837	5557
RXA00230	GR00137	5155	5871
RXA00235	GR00142	1	690
RXA00246	GR00142	641	1054
RXA00247	GR00142	3	506
RXA00247	GR00143	1	2739
RXA00248	GR00143	502	897
RXA00249	GR00143	3744	4148
RXA00250	GR00156	2916	2245
RXA00250	GR00156	2980	3327
RXA00254	GR00145	1606	1138
RXA00254	GR00151	1	11894
RXA00254	GR00151	11577	14582
RXA00257	GR00156	9442	8924
RXA00262	GR00156	11894	14220
RXA00265	GR00156	14220	3
RXA00265	GR00159	3	332

Identification	Code	NT	Stop
	Config	Start	
RXA00597	GR00159	797	1066
RXA00598	GR00159	1070	1387
RXA00601	GR00159	3459	3749
RXA00604	GR00159	5489	5779
RXA00616	GR00162	3574	3918
RXA00617	GR00162	4002	5084
RXA00631	GR00166	172	1626
RXA00646	GR00169	446	6
RXA00647	GR00169	641	1273
RXA00652	GR00169	5449	5997
RXA00653	GR00169	6924	6160
RXA00656	GR00169	9485	9215
RXA00661	GR00172	864	1353
RXA00662	GR00172	2671	1403
RXA00684	GR00173	635	1219
RXA00678	GR00178	647	1393
RXA00679	GR00179	1037	103
RXA00692	GR00181	1450	2317
RXA00693	GR00181	4303	3821
RXA00701	GR00182	427	801
RXA00704	GR00183	2972	3484
RXA00707	GR00185	377	1348
RXA00712	GR00187	1048	500
RXA00714	GR00188	1809	1249
RXA00720	GR00188	7665	7000
RXA00721	GR00189	361	5
RXA00723	GR00190	537	4
RXA00724	GR00191	811	164
RXA00725	GR00191	458	808
RXA00726	GR00192	841	701
RXA00728	GR00194	1	642
RXA00729	GR00194	1063	731
RXA00730	GR00194	1063	731
RXA00739	GR00202	819	4
RXA00740	GR00202	1646	1088
RXA00741	GR00202	2986	2054
RXA00742	GR00202	5517	3888
RXA00743	GR00202	6652	6230
RXA00745	GR00202	13874	13341
RXA00748	GR00202	13755	14945
RXA00747	GR00202	15067	15634
RXA00748	GR00202	15917	16360
RXA00749	GR00202	17240	16542
RXA00750	GR00202	18937	19374
RXA00751	GR00202	20245	19418
RXA00752	GR00202	21847	21419
RXA00754	GR00203	244	664
RXA00757	GR00203	3119	4372
RXA00769	GR00204	6624	6836

Identification	Code	NT	Stop
	Config.	Start	
RXA00771	GR00205	857	100
RXA00785	GR00207	625	5
RXA00786	GR00209	910	686
RXA00785	GR00211	4228	4755
RXA00804	GR00215	438	881
RXA00811	GR00218	1695	2198
RXA00812	GR00219	287	1345
RXA00814	GR00219	2463	3238
RXA00815	GR00219	3236	3808
RXA00816	GR00219	4182	4878
RXA00825	GR00223	587	37
RXA00831	GR00224	1662	961
RXA00838	GR00226	197	2487
RXA00837	GR00227	540	247
RXA00840	GR00229	742	1455
RXA00841	GR00228	1406	2002
RXA00853	GR00231	3775	3173
RXA00854	GR00231	4708	4920
RXA00855	GR00232	526	242
RXA00862	GR00236	580	17
RXA00876	GR00241	4208	2454
RXA00881	GR00242	8057	8434
RXA00882	GR00242	8788	9485
RXA00883	GR00242	10660	9842
RXA00893	GR00244	789	193
RXA00895	GR00244	2578	1988
RXA00904	GR00246	1457	702
RXA00908	GR00247	1611	2168
RXA00914	GR00250	1271	6
RXA00915	GR00251	514	5
RXA00916	GR00251	4108	518
RXA00917	GR00251	5534	4152
RXA00919	GR00252	1890	682
RXA00920	GR00252	2852	1890
RXA00921	GR00252	4750	2852
RXA00922	GR00252	6409	4823
RXA00923	GR00252	6857	6684
RXA00924	GR00252	7278	6817
RXA00925	GR00252	8546	7281
RXA00932	GR00253	5068	5341
RXA00933	GR00253	6047	5586
RXA00943	GR00258	3	509
RXA00946	GR00259	3034	3807
RXA00959	GR00285	402	728
RXA00961	GR00289	442	5
RXA00969	GR00273	1	147
RXA00971	GR00273	1421	1149
RXA00973	GR00274	2272	1670

Identification Code	Contig	Start	Stop	NT
RXA00978	GR00276	217	831	
RXA00988	GR00280	1371	949	
RXA01005	GR00286	520	1365	
RXA01007	GR00287	2572	866	
RXA01008	GR00287	2719	4859	
RXA01016	GR00290	1141	494	
RXA01028	GR00295	3	826	
RXA01029	GR00295	1338	1826	
RXA01031	GR00295	3182	3847	
RXA01032	GR00295	3974	4148	
RXA01033	GR00295	4363	4898	
RXA01034	GR00295	5177	4824	
RXA01035	GR00295	5816	8423	
RXA01036	GR00295	6513	6965	
RXA01037	GR00295	7000	7527	
RXA01038	GR00295	7510	8276	
RXA01039	GR00295	9540	8965	
RXA01040	GR00295	9711	10613	
RXA01041	GR00295	10760	10932	
RXA01042	GR00295	11088	12185	
RXA01043	GR00295	12774	13346	
RXA01044	GR00295	14024	15290	
RXA01045	GR00295	15407	17230	
RXA01046	GR00295	17441	19219	
RXA01047	GR00295	19244	19717	
RXA01058	GR00298	8588	8246	
RXA01062	GR00297	490	5	
RXA01063	GR00297	828	499	
RXA01066	GR00298	605	1330	
RXA01069	GR00299	606	4	
RXA01075	GR00300	3289	2859	
RXA01083	GR00302	1777	1502	
RXA01085	GR00303	960	4	
RXA01086	GR00304	2	463	
RXA01082	GR00305	702	881	
RXA01098	GR00306	4341	3843	
RXA01103	GR00306	10316	10092	
RXA01107	GR00306	13612	14811	
RXA01108	GR00306	15582	14912	
RXA01109	GR00308	16281	15640	
RXA01112	GR00307	1	870	
RXA01121	GR00310	2479	3156	
RXA01122	GR00311	557	36	
RXA01123	GR00311	1090	644	
RXA01127	GR00314	2	280	
RXA01128	GR00314	1325	312	
RXA01131	GR00315	445	1311	
RXA01134	GR00317	2	460	

Identification	Code	NT	Stop
	Config.	Start	
RXA01117	GR00318	1101	1480
RXA01140	GR00318	3272	4057
RXA01148	GR00323	1452	2051
RXA01153	GR00326	546	4
RXA01154	GR00326	808	6
RXA01155	GR00327	1170	6
RXA01156	GR00327	1588	1388
RXA01160	GR00328	4187	3213
RXA01163	GR00331	710	6
RXA01170	GR00334	638	1120
RXA01171	GR00334	1714	2406
RXA01188	GR00332	3005	2523
RXA01187	GR00333	3	323
RXA01169	GR00334	1	567
RXA01176	GR00334	638	
RXA01171	GR00334	1714	
RXA01173	GR00334	4853	4239
RXA01174	GR00334	6004	5255
RXA01178	GR00335	4106	4555
RXA01184	GR00338	1489	17
RXA01187	GR00338	3850	4308
RXA01208	GR00346	593	853
RXA01210	GR00349	3	695
RXA01213	GR00351	1508	282
RXA01218	GR00351	1078	1506
RXA01211	GR00356	1384	1887
RXA01233	GR00356	4242	3871
RXA01234	GR00357	633	250
RXA01256	GR00365	5613	5195
RXA01263	GR00367	10720	11631
RXA01287	GR00387	16799	15486
RXA01275	GR00387	28418	29335
RXA01276	GR00387	29993	30538
RXA01281	GR00389	3869	4630
RXA01295	GR00371	3764	4738
RXA01298	GR00373	5836	4754
RXA01301	GR00375	1993	1589
RXA01304	GR00376	1982	2467
RXA01308	GR00376	5691	4684
RXA01310	GR00380	803	477
RXA01313	GR00381	1116	172
RXA01315	GR00382	1394	744
RXA01316	GR00382	1855	1563
RXA01317	GR00382	2296	1877
RXA01326	GR00386	45	338
RXA01330	GR00387	569	1024
RXA01333	GR00389	1231	227
RXA01336	GR00389	3640	3038
RXA01337	GR00389	5005	3653

Identification	Code	NT	Stop
	Config.	NT	Start
RXA01149	GR00392	1531	755
RXA01176	GR00401	1281	4
RXA01181	GR00108	1147	5
RXA01184	GR00406	J238	1523
RXA01186	GR00408	992	336
RXA01191	GR00408	2078	1J89
RXA01190	GR00409	2988	1489
RXA01401	GR00409	1193	1453
RXA01402	GR00409	3508	1981
RXA01403	GR00409	4410	3997
RXA01405	GR00410	1844	1J89
RXA01410	GR00411	866	4
RXA01413	GR00412	854	1463
RXA01414	GR00412	1628	2134
RXA01415	GR00412	2192	4615
RXA01417	GR00414	845	49
RXA01421	GR00416	1215	829
RXA01422	GR00416	2003	1221
RXA01434	GR00417	10228	9863
RXA01440	GR00418	7498	8489
RXA01441	GR00418	8542	7514
RXA01445	GR00418	15081	14091
RXA01447	GR00418	17885	18733
RXA01448	GR00418	19798	19017
RXA01452	GR00419	2363	2641
RXA01456	GR00420	898	1419
RXA01457	GR00420	1499	2173
RXA01459	GR00420	J111	4120
RXA01460	GR00420	4068	4J59
RXA01469	GR00422	2091	J122
RXA01470	GR00422	4112	3687
RXA01471	GR00422	5243	4437
RXA01472	GR00422	5783	5328
RXA01473	GR00422	6596	5832
RXA01474	GR00422	6878	7223
RXA01475	GR00422	7651	7226
RXA01476	GR00422	7847	8188
RXA01479	GR00422	12423	12650
RXA01484	GR00422	20468	19523
RXA01485	GR00422	20230	22281
RXA01518	GR00424	21238	23711
RXA01519	GR00424	23725	24471
RXA01520	GR00424	24784	25167
RXA01525	GR00424	32301	30580
RXA01527	GR00425	5126	2616
RXA01529	GR00426	2	277
RXA01536	GR00427	4086	2825
RXA01539	GR00428	120	2042

<u>Identification</u>	<u>NT</u>	<u>Stop</u>
<u>Code</u>	<u>Config</u>	<u>Start</u>
RXA01540	GR00428	3083
RXA01542	GR00429	3
RXA01543	GR00430	2802
RXA01544	GR00430	3498
RXA01545	GR00430	4838
RXA01546	GR00430	5584
RXA01547	GR00430	6171
RXA01548	GR00430	7432
RXA01549	GR00430	8428
RXA01552	GR00431	6122
RXA01554	GR00432	3719
RXA01557	GR00433	859
RXA01560	GR00435	767
RXA01565	GR00437	1178
RXA01566	GR00437	1656
RXA01567	GR00437	2213
RXA01574	GR00438	6963
RXA01575	GR00438	8024
RXA01579	GR00439	871
RXA01586	GR00441	1597
RXA01597	GR00442	120
RXA01590	GR00445	1710
RXA01598	GR00447	7414
RXA01602	GR00447	13591
RXA01605	GR00448	960
RXA01610	GR00449	4343
RXA01611	GR00449	4832
RXA01612	GR00449	5235
RXA01618	GR00451	1387
RXA01619	GR00451	2407
RXA01627	GR00453	1
RXA01628	GR00453	8668
RXA01630	GR00454	341
RXA01634	GR00454	4988
RXA01638	GR00456	825
RXA01639	GR00456	1334
RXA01641	GR00456	5182
RXA01642	GR00456	6557
RXA01643	GR00456	6557
RXA01665	GR00463	2152
RXA01672	GR00467	2
RXA01675	GR00467	2824
RXA01676	GR00467	4179
RXA01681	GR00467	10681
RXA01688	GR00470	2026

Identification	Code	NT	Stop
	Config	NT	Start
RXA01694	GR00474	3911	3932
RXA01697	GR00476	761	1486
RXA01701	GR00478	196	528
RXA01703	GR00479	2118	1848
RXA01708	GR00482	312	602
RXA01711	GR00484	2007	850
RXA01714	GR00485	885	371
RXA01729	GR00489	2636	3154
RXA01731	GR00491	109	607
RXA01734	GR00492	544	1077
RXA01741	GR00493	7535	6738
RXA01742	GR00493	7614	8117
RXA01750	GR00498	1878	3516
RXA01751	GR00496	5576	5830
RXA01752	GR00497	557	6
RXA01753	GR00497	2095	557
RXA01754	GR00497	4082	2142
RXA01760	GR00498	5095	5376
RXA01761	GR00499	7001	5484
RXA01765	GR00500	3144	4085
RXA01767	GR00501	341	6
RXA01768	GR00501	827	450
RXA01769	GR00501	1275	847
RXA01770	GR00501	5134	1370
RXA01771	GR00502	886	185
RXA01773	GR00503	34	444
RXA01774	GR00503	634	1416
RXA01775	GR00504	178	741
RXA01776	GR00504	838	2269
RXA01777	GR00504	2319	2777
RXA01778	GR00504	2912	4048
RXA01779	GR00504	4246	5684
RXA01780	GR00504	5721	6095
RXA01781	GR00504	6032	6312
RXA01782	GR00504	6384	6779
RXA01783	GR00504	6882	7078
RXA01785	GR00505	729	1304
RXA01787	GR00506	2	355
RXA01788	GR00506	361	801
RXA01789	GR00506	875	1516
RXA01790	GR00506	1672	1731
RXA01791	GR00506	1885	2247
RXA01792	GR00506	2310	2582
RXA01793	GR00506	2916	3149
RXA01794	GR00506	3194	3427
RXA01799	GR00509	377	1510
RXA01800	GR00509	2292	1573
RXA01809	GR00510	1	638

Identification Code	Config	NT Start	NT Stop
RXA01812	GR00514	3	1232
RXA01813	GR00515	635	6
RXA01816	GR00515	420	4941
RXA01817	GR00515	4941	5573
RXA01820	GR00515	8380	9733
RXA01825	GR00516	2847	2578
RXA01831	GR00516	10814	10411
RXA01834	GR00517	2478	1777
RXA01842	GR00522	1397	480
RXA01843	GR00522	876	1067
RXA01845	GR00522	1919	2326
RXA01846	GR00523	261	4
RXA01847	GR00524	52	786
RXA01854	GR00525	5892	5966
RXA01855	GR00526	1	1838
RXA01858	GR00527	225	770
RXA01857	GR00527	939	1589
RXA01858	GR00529	578	6
RXA01870	GR00534	2123	2797
RXA01874	GR00535	2556	2803
RXA01893	GR00544	1874	2659
RXA01902	GR00544	7957	1094
RXA01903	GR00545	3	281
RXA01904	GR00545	762	340
RXA01905	GR00545	1074	1604
RXA01906	GR00545	2322	2788
RXA01907	GR00545	3176	3787
RXA01908	GR00645	4030	4512
RXA01909	GR00546	59	937
RXA01910	GR00546	1030	1875
RXA01911	GR00546	2189	3044
RXA01921	GR00551	943	5
RXA01923	GR00645	1311	1739
RXA01924	GR00552	1	837
RXA01925	GR00553	1008	1674
RXA01930	GR00546	2189	3044
RXA01931	GR00555	3817	2867
RXA01941	GR00551	943	5
RXA01956	GR00552	1311	1739
RXA01957	GR00553	1	837
RXA01958	GR00564	389	850
RXA01959	GR00564	910	1416
RXA01960	GR00557	995	1429
RXA01961	GR00585	187	504
RXA01962	GR00564	521	1000
RXA01963	GR00565	1022	1591
RXA01963	GR00565	1757	2440
RXA01964	GR00566	1329	4
RXA01965	GR00566	1935	1375
RXA01969	GR00567	5889	5216

Identification	Code	NT	Start	Stop
	Config.	NT	Start	Stop
RXA01973	GR00570	58J		
RXA01974	GR00570	2	658	2109
RXA01976	GR00570		658	2222
RXA01977	GR00571		3742	3972
RXA01978	GR00572		4547	1187
RXA01981	GR00573		1	2105
RXA01987	GR00576		2583	167
RXA01988	GR00576		2384	379
RXA01990	GR00581		928	1720
RXA01991	GR00581		1	999
RXA01999	GR00589		2854	999
RXA02001	GR00589		700	152
RXA02003	GR00593		501	4
RXA02004	GR00594		3	209
RXA02005	GR00594		166	447
RXA02006	GR00597		498	4
RXA02007	GR00598		651	223
RXA02009	GR00601		127	5
RXA02011	GR00603		46	363
RXA02013	GR00607		553	5
RXA02014	GR00607		935	540
RXA02019	GR00612		597	106
RXA02021	GR00613		2008	1061
RXA02038	QR00619		3441	3821
RXA02039	GR00621		3	812
RXA02040	GR00621		1452	925
RXA02045	GR00623		1913	2173
RXA02046	GR00623		2680	2943
RXA02049	GR00624		1583	2039
RXA02050	GR00624		2462	2833
RXA02051	GR00624		3188	3683
RXA02053	GR00624		5484	6082
RXA02058	GR00625		4051	3500
RXA02059	GR00625		4876	4184
RXA02066	GR00626		6187	8678
RXA02067	GR00626		6713	7188
RXA02069	GR00627		1116	1694
RXA02081	GR00628		12307	13905
RXA02084	GR00629		2920	2576
RXA02089	GR00629		8431	8901
RXA02090	GR00629		9764	8984
RXA02091	GR00629		10512	9862
RXA02094	GR00629		13282	13998
RXA02097	GR00630		184	3555
RXA02102	GR00631		4479	3322
RXA02103	GR00631		4510	4905
RXA02109	GR00632		3460	2540
RXA02114	GR00634		815	120

Identification	Code	NT	Stop
Config.	Start	NT	Stop
RXA02121	GR006J6	5813	5109
RXA02125	GR006J7	739	1539
RXA02129	GR006J7	5906	6139
RXA02146	GR006J9	14742	15368
RXA02151	GR006J9	19913	21100
RXA02152	GR006A0	237	638
RXA02163	GR006A0	10072	10824
RXA02164	GR006A0	10824	12398
RXA02165	GR006A0	12388	12999
RXA02168	GR006A1	2894	81
RXA02169	GR006A1	3172	4017
RXA02170	GR006A1	4798	4025
RXA02176	GR006A1	13628	14497
RXA02181	GR006A1	17168	17845
RXA02185	GR006A1	20185	20763
RXA02186	GR006A1	21213	20995
RXA02199	GR006A6	2591	3160
RXA02203	GR006A6	7469	7092
RXA02208	GR006A8	8927	10862
RXA02207	GR006A6	10909	11667
RXA02212	GR006A9	964	467
RXA02221	GR006S1	6720	8081
RXA02226	GR006S1	1059	4
RXA02227	GR006S1	1236	1853
RXA02230	GR006S3	4158	3620
RXA02231	GR006S1	5111	4356
RXA02239	GR006S4	5241	5525
RXA02265	GR006S5	653	1165
RXA02267	GR006S5	2053	1181
RXA02271	GR006S5	5406	5963
RXA02279	GR006S1	1	1404
RXA02280	GR006S8	2	754
RXA02283	GR006S0	2	532
RXA02285	GR006S0	1544	2272
RXA02286	GR006S0	3285	3833
RXA02287	GR006S0	4071	4622
RXA02294	GR006S2	5992	5618
RXA02298	GR006S2	6978	7466
RXA02300	GR006S2	11184	10862
RXA02301	GR006S2	11910	11194
RXA02302	GR006S2	12036	12800
RXA02303	GR006S3	1	720
RXA02304	GR006S3	1613	723
RXA02307	GR006S4	395	6
RXA02325	GR006S8	4314	3445
RXA02330	GR006S0	605	15
RXA02331	GR006S1	396	761
RXA02336	GR006S7	2731	2552

Identification	Code	NT	Stop
		Start	Contig.
RXA02330	GR00673	484	5
RXA02339	GR00674	1	492
RXA02340	GR00674	1214	576
RXA02341	GR00675	415	5
RXA02356	GR00684	761	1756
RXA02358	GR00685	1239	1529
RXA02360	GR00685	3644	6076
RXA02361	GR00685	6160	6810
RXA02362	GR00685	7045	10743
RXA02366	GR00687	254	1581
RXA02368	GR00687	2918	2244
RXA02374	GR00688	1626	2246
RXA02381	GR00691	1792	770
RXA02398	GR00698	2841	4370
RXA02401	GR00699	1391	4491
RXA02406	GR00701	1322	774
RXA02412	GR00703	2043	2522
RXA02415	GR00704	655	170
RXA02417	GR00705	4755	2632
RXA02421	GR00705	7237	6428
RXA02423	GR00706	221	6
RXA02428	GR00707	4585	3452
RXA02433	GR00708	2981	3580
RXA02437	GR00708	1661	2470
RXA02444	GR00709	7838	9113
RXA02454	GR00711	3	815
RXA02457	GR00712	1295	2404
RXA02480	GR00712	5839	5336
RXA02481	GR00712	6252	5845
RXA02484	GR00713	1107	1613
RXA02485	GR00713	2014	1616
RXA02486	GR00714	92	6
RXA02487	GR00714	843	419
RXA02473	GR00715	6864	5974
RXA02475	GR00715	9595	8441
RXA02478	GR00716	1245	10
RXA02483	GR00718	1813	1001
RXA02498	GR00720	11016	11819
RXA02500	GR00720	13460	13558
RXA02505	GR00720	18443	18393
RXA02508	GR00720	19484	18603
RXA02510	GR00721	1983	2618
RXA02519	GR00724	19133	128
RXA02520	GR00724	2222	2905
RXA02534	GR00726	5536	6339
RXA02537	GR00726	8961	9422
RXA02538	GR00726	9422	10933
RXA02546	GR00726	19927	18824

Identification	Code	NT	Stop
	Config.	NT	Start
RXA02552	GR00730	120	
RXA02554	GR00731	1050	427
RXA02555	GR00731	1757	1155
RXA02564	GR00732	2543	3217
RXA02568	GR00735	1363	5
RXA02569	GR00736	82	831
RXA02570	GR00736	837	1478
RXA02576	GR00740	1569	148
RXA02577	GR00740	2463	1579
RXA02591	GR00741	15780	17609
RXA02593	GR00741	18893	18481
RXA02594	GR00741	19077	18754
RXA02606	GR00742	13514	12144
RXA02609	GR00742	16197	16445
RXA02610	GR00742	16452	17378
RXA02619	GR00746	204	1103
RXA02620	GR00746	1192	1845
RXA02624	GR00748	5802	4889
RXA02647	GR00751	4153	4616
RXA02649	GR00752	1284	283
RXA02652	GR00752	2973	3551
RXA02655	GR00752	9313	8330
RXA02662	GR00753	1461	1724
RXA02670	GR00753	10189	10780
RXA02673	GR00753	14030	13398
RXA02678	GR00754	1858	4775
RXA02679	GR00754	5288	5893
RXA02680	GR00754	6392	5109
RXA02681	GR00754	5751	6194
RXA02683	GR00754	7742	7065
RXA02685	GR00754	10058	9402
RXA02686	GR00756	742	287
RXA02712	GR00758	13087	12273
RXA02715	GR00758	15847	15458
RXA02725	GR00760	1478	887
RXA02727	GR00760	6287	5316
RXA02734	GR00762	6514	6897
RXA02736	GR00763	1753	797
RXA02744	GR00763	14480	13657
RXA02753	GR00765	26310	138
RXA02756	GR00766	3851	2981
RXA02757	GR00768	4475	3930
RXA02765	GR00769	3552	2794
RXA02770	GR00772	3	1322
RXA02774	GR00773	3	473
RXA02775	GR00773	744	966
RXA02778	GR00773	1713	1372
RXA02777	GR00773	4628	5732

Identification Code	NT Config.	NT Start	NT Stop
RXA02778	GR00773	1005	1019
RXA02779	GR00773	10617	10895
RXA02780	GR00773	10954	11280
RXA02781	GR00774	1145	155
RXA02782	GR00775	204	875
RXA02783	GR00775	845	1393
RXA02784	GR00775	1751	1936
RXA02785	GR00777	2	808
RXA02793	GR00777	9185	8684
RXA02812	GR00783	2	568
RXA02815	GR00798	3	554
RXA02816	GR00797	2	499
RXA02817	GR00798	403	5
RXA02818	GR00799	611	6
RXA02823	GR00804	275	6
RXA02825	GR00806	565	182
RXA02827	GR00812	428	6
RXA02835	GR00824	299	523
RXA02838	GR00831	1	462
RXA02841	GR00840	283	5
RXA02842	GR00841	356	15
RXA02844	GR00843	247	495
RXA02845	GR00844	2	616
RXA02848	GR00845	578	6
RXA02856	CR10003	459	211
RXA02858	CR10004	1392	387
RXA02862	CR10006	1695	2330
RXA02867	CR10008	610	5
RXA02868	CR10008	2017	1262
RXA02869	CR10009	390	4
RXA02870	CR10011	6	144
RXA02871	GR10011	396	830
RXA02876	GR10016	405	1067
RXA02881	GR10019	94	759
RXA02882	CR10020	2	724
RXA02885	GR10021	1	1536
RXA02886	GR10024	328	754
RXA02889	GR10026	1123	2706
RXA02891	GR10035	3	602
RXA02892	GR10035	1171	668
RXA02898	GR10038	256	5
RXA02905	GR10044	477	4
RXA0194	GR00423	8515	7520
RXA0192	GR00305	702	881
RXA01186	GR00338	3742	2645

TABLE 2: GENES IDENTIFIED FROM GENBANK

GenBank™ Accession No.	Gene Name	Gene Function	Reference
A09073	pppB	Phosphoenol pyruvate carboxylase	Bachmann, B. et al "DNA fragment coding for phosphoenolpyruval carboxylase, recombinant DNA carrying said fragment, strains carrying the recombinant DNA and method for producing L-amino acids using said strains," Patent: EP 0358940-A 3 03/21/90
A45579, A45581, A45582, A45585 A45587		Threonine dehydratase	Mocchetti, B. et al. "Production of L-isocituline by means of recombinant micro-organisms with deregulated threonine dehydratase," Patent: WO 9519442-A 5 07/20/95
AB003132	murC, ftsQ, ftsZ		Kobayashi, M. et al. "Cloning, sequencing, and characterization of the ftsZ gene from coryneform bacteria," <i>Biophys. Res Commun.</i> , 236(2):383-388 (1997).
AB015023	murC; ftsQ		Wachi, M. et al. "A murC gene from Coryneform bacteria," <i>Appl. Microbiol. Biotechnol.</i> , 51(2):223-228 (1999)
AB018330	ftsR		Kimura, E. et al. "Molecular cloning of a novel gene, ftsR, which rescues the detergent sensitivity of a mutant derived from <i>Brevibacterium lacticfermentum</i> ," <i>Biochi. Biotechnol. Biochem.</i> , 60(10):1565-1570 (1996)
AB018331	ftsR1; ftsR2	D-glutamate racemase	
AB020624	murJ	transketolase	
AB023377	ktl	Glutamine 2-oxoglutaric aminotransferase	
AB024708	gltB; gltD	large and small subunits	
AB025224	acn	aconitase	
AB027114	rep	Replication protein	
AB027715	rep; aad	Replication protein, aminoglycoside acetyltransferase	
AF005242	argC	N-acetylglutamate- γ -semialdehyde dehydrogenase	
AF005635	glnA	Glutamine synthetase	
AF030405	hisF	cyclase	
AF030520	argG	Λ glutamino succinate synthetase	
AF031518	argF	Ornithine carbamoyltransferase	
AF036932	ainD	3-dehydroquinic dehydrogenase	

GenBank™ Accession No.	Gene Name	Gene Function	Reference
AF038548	pyc	Pyruvate carboxylase	
AF038651	dcrA; epi; rel	Dipeptide-binding protein; adenine phosphoribosyltransferase; GTP pyrophosphokinase	Wachmeier, L. et al. "The role of the <i>Corynebacterium glutamicum</i> rel gene in (p)ppGpp metabolism." <i>Microbiology</i> , 144, 1853-1862 (1998)
AF041436	argR	Arginine replicon	
AF045908	impA	Inositol monophosphate phosphatase	
AF048764	argH	Argininosuccinate lyase	
AF049897	argC; argJ; argR; argD; argF; argR; argG; argH	N-acetylglutamylphosphate racutase; ornithine acetyltransferase; N-acetylglutamate kinase; acetylornithine transaminase; ornithine carbamoyltransferase; arginine repressor; argininosuccinate synthase; argininosuccinate lyase	
AF050109	inhA	Enoyl-acyl carrier protein reductase	
AF050166	hisG	ATP phosphoribosylformimino- δ -amino-1-phosphoribosyl-4-imidazoleacboxamid isomerase	
AF051846	hisA	Phosphoribosylformimino- δ -amino-1-phosphoribosyl-4-imidazoleacboxamid isomerase	
AF052652	metA	Homoserine O-acetyltransferase	Park, S. et al. "Isolation and analysis of metA, a methionine biosynthetic gene encoding homoserine acetyltransferase in <i>Corynebacterium glutamicum</i> ." <i>Mol Cells</i> , 8(3):286-294 (1998)
AF053071	aroB	Dehydroquinolinate synthetase	
AF060558	hisJ	Glutaminc amidotransferase	
AF066704	hisE	Phosphoribosyl-ATP pyrophosphohydrolase	
AF111233	aroA	S-enolpyruvylshikimate 3-phosphate synthase	
AF116184	panD	L-aspartate- α -ketoglutarate decarboxylase precursor	Dusch, N. et al. "Expression of the <i>Corynebacterium glutamicum</i> panD gene encoding L-aspartate α -ketoglutarate decarboxylase leads to pantothenate overproduction in <i>Escherichia coli</i> ." <i>Appl. Environ. Microbiol.</i> , 65(4):1530-1539 (1999)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
AF124518	aroD; aroE	3-dehydroquinase; shikimate dehydrogenase	
AF124600	aroC; aroK; aroB; proP; proQ	Chorismate synthase; shikimate kinase; 3-dehydroquinase synthase; putative cytoplasmic peptidase	
AF145897	inhA		
AF145898	inhA	Transpox of ectoine, glycine betaine, proline	Peter, H. et al. "Corynebacterium glutamicum is equipped with four secondary carriers for compatible solutes. Identification, sequencing, and characterization of the proline/ectoine uptake system, ProP, and the ectoine/proline/glycine betaine carrier, EcpP." <i>J. Bacteriol.</i> , 180(22):6015-6012 (1998)
AJ001436	cellP		Wehrmann, A. et al. "Different modes of diaminopimelic acid synthetis and their role in cell wall integrity: A study with <i>Corynebacterium glutamicum</i> ." <i>J. Bacteriol.</i> , 180(12):3159-3165 (1998)
AJ004934	dapD	Tetrahydrodipicolinate succinylase (incomplete)	
AJ007732	ppc; secG; ami; ocd; soxA	Phosphoenolpyruvate-carboxylase; ?; high affinity ammonium uptake protein; putative ornithine-cyclodecarboxylase; sarcosine oxidase	Jakoby, M. et al. "Nitrogen regulation in <i>Corynebacterium glutamicum</i> : Isolation of genes involved in biochemical characterization of corresponding proteins." <i>FEMS Microbiol.</i> , 173(2):303-310 (1999)
AJ010319	tsvY; glnB; glnD; sfp; amiP	Involved in cell division; PII protein; uridylyltransferase (uridyl)-removing enzyme; signal recognition particle; low affinity ammonium uptake protein	
AJ132968	cal	Chloramphenicol acetyl transferase	Molenaa, D. et al. "Biochemical and genetic characterization of the membrane-associated malate dehydrogenase (acep10) from <i>Corynebacterium glutamicum</i> ." <i>Eur J Biochem</i> , 254(2):395-403 (1998)
AJ224946	tnqQ	L-malate: quinone oxidoreductase	
AJ238250	ndh	NADH dehydrogenase	Lichinger, T. et al. "Biochemical and biophysical characterization of the cell wall protein of <i>Corynebacterium glutamicum</i> : The channel is formed by a low-molecular mass polypeptide." <i>Biochemistry</i> , 37(43):15024-15032 (1998)
AJ238703	polA		Veres, A. A. et al. "Isolation and characterization of IS31831, a transposable element from <i>Corynebacterium glutamicum</i> ." <i>Mol Microbiol</i> , 11(4):739-746 (1994)
DI7429		Transposable element IS31831	

GenBank® Accession No.	Gene Name	Gene Function	Reference
DB4102	odhA	2-oxoglutarate dehydrogenase	Usuda, Y. et al. "Molecular cloning of the <i>Corynebacterium glutamicum</i> (Brevibacterium lactofermentum AJ12036) odhA gene encoding a novel type of 2-oxoglutarate dehydrogenase," <i>Microbiology</i> , 142:3347-3354 (1996)
E01358	hdh, hk	L-homoserine dehydrogenase; homoserine kinase	Katsunuma, R. et al. "Production of L-threonine and L-isoleucine," Patent: JP 1987232392-A 1 10/12/87
E01359		Upstream of the start codon of homoserine kinase gene	Katsunuma, R. et al. "Production of L-threonine and L-isoleucine," Patent: JP 1987232392-A 2 10/12/87
E01375		Tryptophan operon Leader peptide; anthranilate synthase	Matsui, K. et al. "Tryptophan operon, peptide and protein coded thereby, utilization of tryptophan operon gene expression and production of tryptophan," Patent: JP 1987244362-A 1 10/24/87
E01376	trpL; trpE	Promoter and operator regions of tryptophan operon	Matsui, K. et al. "Tryptophan operon, peptide and protein coded thereby, utilization of tryptophan operon gene expression and production of tryptophan," Patent: JP 1987244382-A 1 10/24/87
E01377		Biotin synthase	Hatakeyama, K. et al. "DNA fragment containing gene capable of coding biotin synthetase and its utilization," Patent: JP 1992278088-A 1 10/02/92
E03937		Diamino pelargonate acid aminotransferase	Kohama, K. et al. "Gene coding diaminopelargonate acid aminotransferase and deshionoin synthetase and its utilization," Patent: JP 19922330284-A 1 11/18/92
E04040		Deshionoin synthetase	Kohama, K. et al. "Gene coding diaminopelargonate acid aminotransferase and deshionoin synthetase and its utilization," Patent: JP 19922330284-A 1 11/18/92
E04041			Kurusu, Y. et al. "Gene DNA coding aspartase and utilization thereof," Patent: JP 1993030977-A 1 02/09/93
E04307		Flavum aspartase	Katsunuma, R. et al. "Gene manifestation controlling DNA," Patent: JP 1993036782-A 3 03/09/93
E04376		Isocitric acid lyase	Katsunuma, R. et al. "Gene manifestation controlling DNA," Patent: JP 1993036782-A 3 03/09/93
E04377		Isocitric acid lyase N-terminal fragment	Solouchi, N. et al. "Production of L-phenylalanine by fermentation," Patent: JP 1993076352-A 2 03/30/93
E04484		Phenylate dehydratase	Fujimoto, N. et al. "Gene DNA coding Aspartokinase and its use," Patent: JP 1993184366-A 1 07/27/93
E05108		Aspartokinase	Itakakeyama, K. et al. "Gene DNA coding dihydrodipicolinic acid synthetase and its use," Patent: JP 1993184371-A 1 07/27/93
E05112		Dihydro-dipiclorimic synthetase	

Table 2. Page 4

GenBank™ Accession No.	Gene Name	Gene Function	Reference
E05776	Diaminopimelic acid dehydrogenase	Kobayashi, M. et al. "Gene DNA coding Diaminopimelic acid dehydrogenase and its use," Patent. JP 1993284970-A 1 11/02/93	
E05779	Threonine Synthase	Kohama, K. et al. "Gene DNA coding threonine synthase and its use," Patent. JP 1993284972-A 1 11/02/93	
E06110	Prephenate dehydratase	Kikuchi, T. et al. "Production of L-phenylalanine by fermentation method," Patent. JP 1993344881-A 1 12/27/93	
E06111	Mutated Prephenate dehydratase	Kikuchi, T. et al. "Production of L-phenylalanine by fermentation method," Patent. JP 1993344881-A 1 12/27/93	
E06146	Acetohydroxy acid synthetase	Inui, M. et al. "Gene capable of coding Acetohydroxy acid synthetase and its use," Patent. JP 1993344891-A 1 12/27/93	
E066825	Aspartokinase	Sugimoto, M. et al. "Mutant aspartokinase gene," patent. JP 1994062866-A 1 03/08/94	
E06826	Mutated aspartokinase alpha subunit	Sugimoto, M. et al. "Mutant aspartokinase gene," patent. JP 1994062866-A 1 03/08/94	
E06827	Mutated aspartokinase alpha subunit	Sugimoto, M. et al. "Mutant aspartokinase gene," patent. JP 1994062866-A 1 03/08/94	
E07701	sccY	Honno, N. et al. "Gene DNA participating in integration of membrane protein to membrane," Patent. JP 1994169780-A 1 06/21/94	
E08177	Aspartokinase	Sato, Y. et al. "Genetic DNA capable of coding Aspartokinase released from feedback inhibition and its utilization," Patent. JP 1994261766-A 1 09/20/94	
E08178,		Sato, Y. et al. "Genetic DNA capable of coding Aspartokinase released from feedback inhibition and its utilization," Patent. JP 1994261766-A 1 09/20/94	
E08179,			
E08180,			
E08181,			
E08182			
E08232	Acetohydroxy acid isomeroeductase	Inui, M. et al. "Gene DNA coding acetohydroxy acid isomeroeductase," Patent. JP 1994277067-A 1 10/04/94	
E08234	sccE	Asai, Y. et al. "Gene DNA coding for translocation machinery of proein," Patent. JP 1994277073-A 1 10/04/94	
E08643	FT aminotransferase and desulphobiotin synthetase promoter region	Itaiakeyama, K. et al. "DNA fragment having promoter function in coryneform bacterium," Patent. JP 1995031476-A 1 02/03/95	
E08646	Biotin synthetase	Itaiakeyama, K. et al. "DNA fragment having promoter function in coryneform bacterium," Patent. JP 1995031476-A 1 02/03/95	

GenBank™ Accession No.	Gene Name	Gene Function	Reference
E08649		Aspartase	Kohama, K. et al "DNA fragment having promoter function in coryneform bacterium," Patent. JP 1995031478-A 102/03/95
E08910		Dihydronicolinate reductase	Medoni, M. et al "DNA fragment containing gene coding Dihydronicolinate acid reductase and utilization thereof," Patent. JP 1995075578-A 103/20/95
E08901		Diaminopimelic acid decarboxylase	Madoi, M. et al. "DNA fragment containing gene coding Diaminopimelic acid decarboxylase and utilization thereof," Patent. JP 1995075579-A 103/20/95
E12394		Seinc hydroxymethylhydrolase transposase	Halakcyama, K. et al. "Production of L-hydrophan," Patent. JP 1997028391-A 102/04/97
E12760, E12759, E12758			Moriya, M. et al "Amplification of gene using artificial transposon," Patent. JP 1997070291-A 03/18/97
E12764		Arginyl-tRNA synthetase; diaminopimelic acid decarboxylase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent. JP 1997070291-A 03/18/97
E12767		Dihydronicolinic acid synthetase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent. JP 1997070291-A 03/18/97
E12770		aspartokinase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent. JP 1997070291-A 03/18/97
E12773		Dihydronicolinic acid reductase	Moriya, M. et al. "Amplification of gene using artificial transposon," Patent. JP 1997070291-A 03/18/97
E13655		Glucose-6-phosphate dehydrogenase	Halakcyama, K. et al. "Glucose-6-phosphate dehydrogenase and DNA capable of coding the same," Patent. JP 1997224661-A 109/02/97
L01508	IlvA	Thiokinase dehydratase	Morckel, B. et al. "Functional and structural analysis of the thiokinase dehydratase of Corynebacterium glutamicum," <i>J. Bacteriol.</i> , 174:8065-8072 (1992)
L07603	EC 4.2.1.15	3-deoxy-D-arabinohexulonate-7-phosphate synthase	Chen, C. et al. "The cloning and nucleotide sequence of <i>Corynebacterium glutamicum</i> 3-deoxy-D-arabinohexulosonate-7-phosphate synthase gene," <i>FEBS Microbiol Lett.</i> , 107:223-230 (1993)
L09232	IlvB; ilvN; ilvC	Acetohydroxy acid synthase large subunit; Acetohydroxy acid synthase small subunit; Acetohydroxy acid isomerase	Keilhauer, C. et al. "Isolation, synthesis in <i>Corynebacterium glutamicum</i> , molecular analysis of the ilvB-ilvN-ilvC operon," <i>J. Bacteriol.</i> , 175(17):5593-5603 (1993)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
L18874	psm	Phosphoenolpyruvate sugar phosphotransferase	Fourel, A et al. "Bacillus subtilis sucrose-specific enzyme II of the phosphotransferase system: expression in <i>Escherichia coli</i> and homology to enzymes II from enteric bacteria," <i>PNAS USA</i> , 84(24) 8773-8777 (1987); Lee, J.K. et al. "Nucleotide sequence of the gene encoding the <i>Corynebacterium glutamicum</i> mannose enzyme II and analyses of the deduced protein sequence," <i>FEMS Microbiol. Lett.</i> , 119(1-2), 137-145 (1994)
L27123	aceB	Malate synthase	Lee, H-S. et al. "Molecular characterization of aceB, a gene encoding malate synthase in <i>Corynebacterium glutamicum</i> ," <i>J. Microbiol. Biotechnol.</i> , 4(4) 256-263 (1994)
L27126		Pyruvate kinase	Jelten, M. S. et al. "Structural and functional analysis of pyruvate kinase from <i>Corynebacterium glutamicum</i> ," <i>Appl. Environ. Microbiol.</i> , 60(7):2501-2507 (1994)
L28760	aceV	Isocitrate lyase	Oguiza, J.A. et al. "Molecular cloning, DNA sequence analysis, and characterization of the <i>Corynebacterium diphtheriae</i> dtxR from <i>Brevibacterium laevigatum</i> ," <i>J. Bacteriol.</i> , 177(2):465-467 (1995)
L35906	dtxI	Diphtheria toxin repressor	Follette, M.T. et al. "Molecular cloning and nucleotide sequence of the <i>Corynebacterium glutamicum</i> phcA gene," <i>J. Bacteriol.</i> , 167:695-702 (1986)
M13774		Peptidate dehydratase	Paik, Y-H. et al. "Phylogenetic analysis of the coryneform bacteria by 5S rRNA sequences," <i>J. Bacteriol.</i> , 169:1801-1806 (1987)
M16175	5S rRNA		Sano, K. et al. "Structure and function of the trp operon control regions of <i>Brevibacterium laevigatum</i> , a glutamic-acid-producing bacterium," <i>Gene</i> , 52:191-200 (1987)
M16663	trpE	Anthranilate synthase, 5' end	Sano, K. et al. "Structure and function of the trp operon control regions of <i>Brevibacterium laevigatum</i> , a glutamic-acid-producing bacterium," <i>Gene</i> , 52:191-200 (1987)
M16664	trpA	Tryptophan synthase, 3' end	O'Regan, M. et al. "Cloning and nucleotide sequence of the phosphoenolpyruvate carboxylase gene of <i>Corynebacterium glutamicum</i> ATCC13032," <i>Gene</i> , 77(2):237-251 (1989)
M23819		Phosphoenolpyruvate carboxylase	Rollet, C. et al. "Gram-positive bacteria with a high DNA G+C content are characterized by a common insertion within their 23S rRNA genes," <i>J. Gen. Microbiol.</i> , 138:1167-1175 (1992)
M85106		23S rRNA gene insertion sequence	

GenBank™ Accession No.	Gene Name	Gene Function	Reference
M85107, M85108		23S rRNA gene insertion sequence	Roller, C. et al. "Gram-positive bacteria with a high DNA G+C content are characterized by a common insertion within their 23S rRNA genes," <i>J. Gen. Microbiol.</i> , 138, 1167-1175 (1992)
M89931	aceD; bmrQ; yhhw	Beta C-S lyase, branched chain amino acid uptake carrier, hypothetical protein yhhw	Rossol, I. et al. "The <i>Corynebacterium glutamicum aceD</i> gene encodes a C-S lyase with alpha, beta-elimination activity that degrades aminoethylcysteine," <i>J. Bacteriol.</i> , 174(9):2968-2977 (1992); Tauch, A. et al. "Isoleucine uptake in <i>Corynebacterium glutamicum</i> ATCC 13032 is directed by the <i>bmrQ</i> gene product," <i>Arch. Microbiol.</i> , 169(4):303-312 (1998)
S59299	trp	LacZα gene (promoter)	Henry, D.M. et al. "Cloning of the <i>trp</i> gene cluster from a tryptophan hyperproducing strain of <i>Corynebacterium glutamicum</i> : identification of a mutation in the <i>trp</i> leader sequence," <i>Appl. Environ. Microbiol.</i> , 59(3):791-799 (1993)
U11545	trpD	Anthranilate phosphoribosyltransferase	O'Gara, J.P. and Dunican, I.K. (1994) Complete nucleotide sequence of the <i>Corynebacterium glutamicum</i> ATCC 21850 <i>trpD</i> gene. "Thesis, Microbiology Department, University College Galway, Ireland.
U13922	cglM; cglR; cglT	Putative type II 5'-cytosine methyltransferase; putative type II restriction endonuclease; putative type I or type III restriction endonuclease	Schaefer, A. et al. "Cloning and characterization of a DNA region encoding a stress-sensitive restriction system from <i>Corynebacterium glutamicum</i> ATCC 13032 and analysis of its role in intergeneric conjugation with <i>Escherichia coli</i> ," <i>J. Bacteriol.</i> , 176(23):7309-7319 (1994); Schaefer, A. et al. "The <i>Corynebacterium glutamicum</i> cglM gene encoding a 5'-cytosine in an <i>MetC</i> -deficient <i>Escherichia coli</i> strain," <i>Gene</i> , 203(2):95-101 (1997)
U14965	recA		Ankri, S. et al. "Mutations in the <i>Corynebacterium glutamicum</i> proline biosynthetic pathway: A natural bypass of the proA step," <i>J. Bacteriol.</i> , 178(15):4412-4419 (1996)
U31224	ppx		Ankri, S. et al. "Mutations in the <i>Corynebacterium glutamicum</i> proline biosynthetic pathway: A natural bypass of the proA step," <i>J. Bacteriol.</i> , 178(15):4412-4419 (1996)
U31225	proC	L-proline NADP ⁺ 5-oxidoreductase	Ankri, S. et al. "Mutations in the <i>Corynebacterium glutamicum</i> proline biosynthetic pathway: A natural bypass of the proA step," <i>J. Bacteriol.</i> , 178(15):4412-4419 (1996)
U31230	obj; proB; unkdh	γ-gamma glutamyl kinase; similar to D-isomer specific 2-hydroxyacid dehydrogenases	Ankri, S. et al. "Mutations in the <i>Corynebacterium glutamicum</i> proline biosynthetic pathway: A natural bypass of the proA step," <i>J. Bacteriol.</i> , 178(15):4412-4419 (1996)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
U31281	bioB	Biotin synthase	Serebrenikin, I.G., "Two new members of the bio B superfamily: Cloning, sequencing and expression of bio B genes of <i>Methylbacterillus flagellatum</i> and <i>Corynebacterium glutamicum</i> ," <i>Gene</i> , 175:15-22 (1996)
U35073	mrR, accBC	Thiosulfate sulfurtransferase; acyl CoA carboxylase	Jager, W. et al. "A <i>Corynebacterium glutamicum</i> gene encoding a two-domain protein similar to bioin carboxylases and biotin-carboxyl-carrier proteins," <i>Arch Microbiol.</i> , 166(2):76-82 (1996)
U43535	cmr	Multidrug resistance protein	Jager, W. et al. "A <i>Corynebacterium glutamicum</i> gene conferring multidrug resistance in the heterologous host <i>Escherichia coli</i> ," <i>J Bacteriol.</i> , 179(7):2449-2451 (1997)
U43536	cprB	Heat shock ATP-binding protein	
U53587	apnA-3	3'S'-aminoglycoside phosphotransferase	
U89648		<i>Corynebacterium glutamicum</i> unidentified sequence involved in histidine biosynthesis, partial sequence	Matsuji, K. et al. "Complete nucleotide and deduced amino acid sequences of the <i>Brevibacterium lactofermentum</i> tryptophan operon," <i>Nucleic Acids Res.</i> , 14(24):10113-10114 (1986)
X04960	trpA; trpB; trpC; trpD; trpE; trpG; trpL	Tryptophan operon	Yeh, P. et al. "Nucleic sequence of the <i>lysA</i> gene of <i>Corynebacterium glutamicum</i> and possible mechanisms for modulation of its expression," <i>Mol Gen Genet</i> , 212(1):112-119 (1988)
X07563	lysA	DAP decarboxylase (meso-diaminopimelate decarboxylase, EC 4.1.1.20)	Eikmanns, B.J. et al. "The Phosphoenolpyruvate carboxylase gene of <i>Corynebacterium glutamicum</i> : Molecular cloning, nucleotide sequence, and expression," <i>Mol Gen Genet</i> , 218(2):330-339 (1989); Lepiniec, L. et al. "Sorghum Phosphoenolpyruvate carboxylase gene family: structure, function and molecular evolution," <i>Plant Mol Biol</i> , 21 (3):487-502 (1993)
X14734	EC 4.1.1.31	Phosphoenolpyruvate carboxylase	Von der Osten, C.H. et al. "Molecular cloning, nucleotide sequence and fine-structural analysis of the <i>Corynebacterium glutamicum</i> (da gene) structural comparison of C. glutamicum fumate-1,6-biphosphate aldolase to class I and class II aldolases," <i>Mol Microbiol</i> .
X17713	ida	Fructose-bisphosphate aldolase	Bonmassi, S. et al. "Nucleic sequence of the <i>dapA</i> gene from <i>Corynebacterium glutamicum</i> ," <i>Nucleic Acids Res.</i> , 18(2):6421 (1990)
X33993	dppA	1,2,3-dihydroxyproline synthetase (EC 4.2.1.52)	

GeoRank TM Accession No.	Gene Name	Gene Function	Reference
X54223		AttB-related site	Cianciotto, N. et al. "DNA sequence homology between att B-related sites of <i>Corynebacterium diphtheriae</i> , <i>Corynebacterium ulcerans</i> , <i>Corynebacterium glutamicum</i> , and the attP site of <i>lambda</i> corynephage," <i>FEMS Microbiol. Lett.</i> , 66:299-302 (1990)
X54740	argS; lysA	Arginyl-tRNA synthetase; Diaminopimelate decarboxylase	Marcel, T. et al. "Nucleotide sequence and organization of the upstream region of the <i>Corynebacterium glutamicum</i> lysA gene," <i>Mol. Microbiol.</i> , 4(11):1819-1830 (1990)
X55994	trpL; trpE	Putative leader peptide; anthranilate synthase component I	Heery, D. M. et al. "Nucleotide sequence of the <i>Corynebacterium glutamicum</i> trpE gene," <i>Nucleic Acids Res.</i> , 18(23):7138 (1990)
X56037	thrC	Threonine synthase	Han, K. S. et al. "The molecular structure of the <i>Corynebacterium glutamicum</i> threonine synthase gene," <i>Mol. Microbiol.</i> , 4(10):1693-1702 (1990)
X56075	attB-related site	Attachment site	Cianciotto, N. et al. "DNA sequence homology between att B-related sites of <i>Corynebacterium diphtheriae</i> , <i>Corynebacterium ulcerans</i> , <i>Corynebacterium glutamicum</i> , and the attP site of <i>lambda</i> corynephage," <i>FEMS Microbiol. Lett.</i> , 66:299-302 (1990)
X57226	lysC-alpha; lysC-beta; asd	Aspartokinase-alpha subunit, Aspartokinase-beta subunit; aspartate beta semialdehyde dehydrogenase	Kalinowski, J. et al. "Genetic and biochemical analysis of the Aspartokinase from <i>Corynebacterium glutamicum</i> ," <i>Mol. Microbiol.</i> , 5(5):1197-1204 (1991); Kalinowski, J. et al. "Aspartokinase genes lysC alpha and lysC beta overlap and are adjacent to the aspartate beta-semialdehyde dehydrogenase gene asd in <i>Corynebacterium glutamicum</i> ," <i>Mol. Gen. Genet.</i> , 224(3):317-324 (1990)
X59403	gap; pgk; fpm	Glyceraldehyde-3-phosphate phosphoglycerate kinase, triosephosphate isomerase	Eikmanns, B.J. "Identification, sequence analysis, and expression of a <i>Corynebacterium glutamicum</i> gene cluster encoding the three glycolytic enzymes glyceraldehyde-3-phosphate dehydrogenase, 3-phosphoglycerate kinase, and triosephosphate isomerase," <i>J. Bacteriol.</i> , 174(19):6076-6086 (1992)
X59404	gdh	Glutamate dehydrogenase	Bormann, E.R. et al. "Molecular analysis of the <i>Corynebacterium glutamicum</i> gdh gene encoding glutamate dehydrogenase," <i>Mol. Microbiol.</i> , 6(3):317-326 (1992)
X60312	lysI	1-lysine permease	Seep-Feldhaus, A. et al. "Molecular analysis of the <i>Corynebacterium glutamicum</i> lysI gene involved in lysine uptake," <i>Mol. Microbiol.</i> , 5(12):2995-3005 (1991)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X66078	cop1	PstI protein	Joliff, G. et al. "Cloning and nucleotide sequence of the <i>csp1</i> gene encoding PS1, one of the two major secreted proteins of <i>Corynebacterium glutamicum</i> : The deduced N-terminal region of PS1 is similar to the Mycobacterium antigen 85 complex," <i>Mol. Microbiol.</i> , 6(16):2349-2362 (1992)
X66112	gltI	Citrate synthase	Eikmanns, B. et al. "Cloning sequence, expression and transcriptional analysis of the <i>Corynebacterium glutamicum gltA</i> gene encoding citrate synthase," <i>Microbiol.</i> , 140:1817-1828 (1994)
X67737	dapB	Dihydrodipicolinate reductase	Peyret, J.L. et al. "Characterization of the <i>csp2</i> gene encoding PS2, an ordered surface-layer protein in <i>Corynebacterium glutamicum</i> ," <i>Mol. Microbiol.</i> , 9(1):97-109 (1993)
X69103	csp2	Surface layer protein PS2	
X69104		IS3 related insertion element	Bonamy, C. et al. "Identification of IS1206, a <i>Corynebacterium glutamicum</i> IS3-related insertion sequence and phylogenetic analysis," <i>Mol. Microbiol.</i> , 14(3):571-581 (1994)
X70959	leuV	Isopropylmalate synthase	Patek, M. et al. "Leucine synthesis in <i>Corynebacterium glutamicum</i> : enzyme activities, structure of <i>leuA</i> , and effect of <i>leuA</i> inactivation on lysine synthesis," <i>Appl. Environ. Microbiol.</i> , 60(1):133-140 (1994)
X71489	icd	Isocitrate dehydrogenase (NADP ⁺)	Eikmanns, B.J. et al. "Cloning sequence analysis, expression, and inactivation of the <i>Corynebacterium glutamicum icd</i> gene encoding isocitrate dehydrogenase and biochemical characterization of the enzyme," <i>J. Bacteriol.</i> , 177(3):774-782 (1995)
X72855	GDH λ	Glutamate dehydrogenase (NADP ⁺)	Heery, D.M. et al. "A sequence from a tryptophan-hyperproducing strain of <i>Corynebacterium glutamicum</i> encoding resistance to 5-methyltryptophan," <i>Biochem. Biophys. Res. Commun.</i> , 201(3):1255-1262 (1994)
X75083	mttA	S-methyltryptophan resistance	
X75084			Fitzpatrick, R. et al. "Construction and characterization of <i>recA</i> mutant strains of <i>Corynebacterium glutamicum</i> and <i>Brevibacterium lactofermentum</i> ," <i>Appl. Microbiol. Biotechnol.</i> , 42(4):575-580 (1994)
X75085	iccA		Reinschmidt, D.J. et al. "Characterization of the isocitrate lyase gene from <i>Corynebacterium glutamicum</i> and biochemical analysis of the enzyme," <i>J. Bacteriol.</i> , 176(12):3474-3483 (1994)
X75504	aceA; thiX	Partial isocitrate lyase; ?	Ludwig, W. et al. "Phylogenetic relationships of bacteria based on comparative sequence analysis of elongation factor Tu and ATP-synthase beta-subunit genes," <i>Antonie Van Leeuwenhoek</i> , 64:285-305 (1993)
X76875		ATPase beta-subunit	

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X77034	wf	Elongation factor Tu	Ludwig, W. et al. "Phylogenetic relationships of bacteria based on comparative sequence analysis of elongation factor Tu and λ TP-synthase beta-subunit genes," <i>Antonie Van Leeuwenhoek</i> , 64:285-305 (1993)
X77384	iceA		Billman-Jacobe, H. "Nucleotide sequence of a <i>iceA</i> gene from <i>Corynebacterium glutamicum</i> ," <i>DNA Seq.</i> , 4(6):403-404 (1994)
X78491	aceB	Malate synthase	Reinscheid, D.J. et al. "Malate synthase from <i>Corynebacterium glutamicum</i> pia-ack operon encoding phosphotransacetylase: sequence analysis," <i>Microbiology</i> , 140:309-3108 (1994)
X80629	16S rDNA	16S ribosomal RNA	Rainey, F.A. et al. "Phylogenetic analysis of the genus <i>Rhodococcus</i> and <i>Nocardioides</i> and evidence for the evolutionary origin of the genus <i>Nocardioides</i> from within the radiation of <i>Rhodococcus</i> species," <i>Microbiol.</i> , 141:523-528 (1995)
X81191	gluA; gluB; gluC; gluD	Glutamate uptake system	Kronemeyer, W. et al. "Structure of the GluABCD cluster encoding the glutamate uptake system of <i>Corynebacterium glutamicum</i> ," <i>J Bacteriol.</i> , 177(5):1152-1158 (1995)
X81379	dapE	Succinylaminopeptidase desuccinylase	Wehrmann, A. et al. "Analysis of different DNA fragments of <i>Corynebacterium glutamicum</i> complementing <i>dapE</i> of <i>Escherichia coli</i> ," <i>Microbiol.</i> , 40:3349-56 (1994)
X82061	16S rDNA	16S ribosomal RNA	Ruimy, R. et al. "Phylogeny of the genus <i>Corynebacterium</i> deduced from analyses of small subunit ribosomal DNA sequences," <i>Int. J. Syst. Bacteriol.</i> , 45(4):740-746 (1995)
X82928	asd; lysC	Aspartate-semialdehyde dehydrogenase; ?	Serebrijski, I. et al. "Muticopy suppression by <i>asd</i> gene and osmotic stress-dependent complementation by heterologous <i>proA</i> in <i>proA</i> mutants," <i>J. Bacteriol.</i> , 177(24):7255-7260 (1995)
X82929	proA	Gamma-glutamyl phosphate reductase	Serebrijski, I. et al. "Muticopy suppression by <i>asd</i> gene and osmotic stress-dependent complementation by heterologous <i>proA</i> in <i>proA</i> mutants," <i>J. Bacteriol.</i> , 177(24):7255-7260 (1995)
X84237	16S rDNA	16S ribosomal RNA	Pascual, C. et al. "Phylogenetic analysis of the genus <i>Corynebacterium</i> based on 16S rRNA gene sequences," <i>Int. J. Syst. Bacteriol.</i> , 45(4):724-728 (1995)
X85965	atop; dapE	Aromatic amino acid permease; ?	Wehrmann, A. et al. "Functional analysis of sequences adjacent to <i>dapE</i> of <i>Corynebacterium glutamicum</i> reveals the presence of <i>atop</i> , which encodes the aromatic amino acid transporter," <i>J Bacteriol.</i> , 177(20):5991-5993 (1995)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X86157	argB, argC; argD; argF; argJ	N-acetyl gamma-glutamyl phosphate reductase; acetylornithine aminotransferase; ornithine carbamoyltransferase; glutamate N-acetyltransferase	Sakanyan, V. et al. "Genes and enzymes of the acetyl cycle of arginine biosynthesis in <i>Corynebacterium glutamicum</i> : enzyme evolution in the early steps of the arginine pathway." <i>Microbiology</i> , 142:99-108 (1996)
X89084	ackA	Phosphate acetyltransferase, acetate kinase	Reinschmidt, D.J. et al. "Cloning, sequence analysis, expression and inactivation of the <i>Corynebacterium glutamicum</i> pha-ack operon encoding phosphotransacetylase and acetate kinase." <i>Microbiology</i> , 145:503-513 (1999)
X89850	attB	Attachment site	Le Marric, C. et al. "Genetic characterization of site-specific integration functions of phi A112 infecting "Arthrobacter aureus C70." <i>J. Bacteriol.</i> 178(7):1996-2004 (1996)
X90356		Promoter fragment F1	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif." <i>Microbiology</i> , 142:1297-1309 (1996)
X90357		Promoter fragment F2	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif." <i>Microbiology</i> , 142:1297-1309 (1996)
X90358		Promoter fragment F10	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif." <i>Microbiology</i> , 142:1297-1309 (1996)
X90359		Promoter fragment F13	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif." <i>Microbiology</i> , 142:1297-1309 (1996)
X90360		Promoter fragment F22	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif." <i>Microbiology</i> , 142:1297-1309 (1996)
X90361		Promoter fragment F34	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif." <i>Microbiology</i> , 142:1297-1309 (1996)
X90362		Promoter fragment F37	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif." <i>Microbiology</i> , 142:1297-1309 (1996)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X90363	Promoter fragment F43	Promoter analysis and search for a consensus motif, "Microbiology"; 142:1297-1309 (1996)	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> ; 142:1297-1309 (1996)
X90364	Promoter fragment F64	Promoter analysis and search for a consensus motif, "Microbiology"; 142:1297-1309 (1996)	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> ; 142:1297-1309 (1996)
X90365	Promoter fragment F75	Promoter analysis and search for a consensus motif, "Microbiology"; 142:1297-1309 (1996)	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> ; 142:1297-1309 (1996)
X90366	Promoter fragment PF101	Promoter analysis and search for a consensus motif, "Microbiology"; 142:1297-1309 (1996)	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> ; 142:1297-1309 (1996)
X90367	Promoter fragment PF104	Promoter analysis and search for a consensus motif, "Microbiology"; 142:1297-1309 (1996)	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> ; 142:1297-1309 (1996)
X90368	Promoter fragment PF109	Promoter analysis and search for a consensus motif, "Microbiology"; 142:1297-1309 (1996)	Patek, M. et al. "Promoters from <i>Corynebacterium glutamicum</i> : cloning, molecular analysis and search for a consensus motif," <i>Microbiology</i> ; 142:1297-1309 (1996)
X93513	amt	Ammonium transport system	Siewe, R. M. et al. "Functional and genetic characterization of the (methyl) ammonium uptake carrier of <i>Corynebacterium glutamicum</i> ," <i>J. Biol. Chem.</i> 271(10):5398-5403 (1996)
X93514	bcpP	Glycine betaine transport system	Petei, I. et al. "Isolation, characterization, and expression of the <i>Corynebacterium glutamicum bcpP</i> gene, encoding the transport system for the compatible solute glycine betaine," <i>J. Bacteriol.</i> 178(17):5229-5234 (1996)
X93649	lysE		Patek, M. et al. "Identification and transcriptional analysis of the dapB-ORF2-dapA-ORF4 operon of <i>Corynebacterium glutamicum</i> , encoding two enzymes involved in L-lysine synthesis," <i>Biochemical Lett.</i> 19:1113-1117 (1997)
X96471	lysE, lysG	Lysine exporter protein, Lysine export regulator protein	Vrijic, M. et al. "A new type of transporter with a new type of cellular function: L-lysine export from <i>Corynebacterium glutamicum</i> ," <i>Mol Microbiol.</i> 27(5):815-826 (1996)

GenBank™ Accession No.	Gene Name	Gene Function	Reference
X96580	panB, panC; xylB	3-methyl-2-oxobutanate hydroxymethyltransferase; pantoate-beta-alanide ligase; xylulokinase	Sahm, H. et al. "D-pantothenate synthesis in <i>Corynebacterium glutamicum</i> and use of panBC and genes encoding L-valine synthesis for D-pantothenate overproduction," <i>Appl. Environ. Microbiol.</i> , 65(5):1973-1979 (1999)
X96962		Insertion sequence IS207 and transposase	Ramos, A. et al. "Cloning, sequencing and expression of the gene encoding elongation factor P in the amino-acid producer <i>Brevibacterium lactofermentum</i> (Corynebacterium glutamicum ATCC 13869)," <i>Gene</i> , 198:217-222 (1997)
X99289		Elongation factor P	Malcos, J., M. et al. "Nucleotide sequence of the homoserine kinase (thiB) gene of the <i>Brevibacterium lactofermentum</i> ," <i>Nucleic Acids Res.</i> , 15(9):1922 (1987)
Y00140	thiB	Homoserine kinase	Ishino, S. et al. "Nucleotide sequence of the meso-diaminopimelate D-dehydrogenase gene from <i>Corynebacterium glutamicum</i> ," <i>Nucleic Acids Res.</i> , 15(9):1917 (1987)
Y00151	ddh	Meso-diaminopimelate D-dehydrogenase (EC 1.4.1.16)	Malcos, L. M. et al. "Nucleotide sequence of the homoserine dehydrogenase (thiA) gene of the <i>Brevibacterium lactofermentum</i> ," <i>Nucleic Acids Res.</i> , 15(24):10598 (1987)
Y00476	thiA	Homoserine dehydrogenase	Peoples, O.P. et al. "Nucleotide sequence and fine structural analysis of the <i>Corynebacterium glutamicum</i> hom-thiB operon," <i> Mol. Microbiol.</i> , 2(1):63-72 (1988)
Y00546	hom; thiB	Homoserine dehydrogenase; homoserine kinase	Homuthia, M. P. et al. "Identification, characterization, and chromosomal organization of the <i>thiZ</i> gene from <i>Brevibacterium lactofermentum</i> ," <i>Mol. Gen. Genet.</i> , 259(1):97-104 (1998)
Y08964	murC, ftsQ/divD; ftsZ	UDP-N-acetylmuramate-alanine ligase, division initiation protein or cell division protein; cell division protein	Peter, H. et al. "Isolation of the <i>ftsP</i> gene of <i>Corynebacterium glutamicum</i> and characterization of a low-affinity uptake system for compatible solutes," <i>Arch. Microbiol.</i> , 168(2):143-151 (1997)
Y09163	ftsP	High affinity proline transport system	Peters-Wendisch, P.G. et al. "Pyruvate carboxylase from <i>Corynebacterium glutamicum</i> : characterization, expression and inactivation of the <i>pyc</i> gene," <i>Microbiology</i> , 144:915-927 (1998)
Y09548	pyc	Pyruvate carboxylase	Paick, M. et al. "Analysis of the <i>leuB</i> gene from <i>Corynebacterium glutamicum</i> ," <i>Appl. Microbiol. Biotechnol.</i> , 50(1):42-47 (1998)
Y105578	leuB	3-isopropylmalic dehydrogenase	Moreau, S. et al. "Site-specific integration of corynephage Phi-16: The construction of an integration vector," <i>Microbiol.</i> , 145:539-548 (1999)
Y12472		Attachment site bacteriophage Phi-16	

Table 2, Page 15

GenBank™ Accession No.	Gene Name	Gene Function	Reference
Y12537	prop	Proline/ecdione uptake system protein	Peter, H. et al. "Corynebacterium glutamicum is equipped with four secondary carriers for compatible solutes. Identification, sequencing, and characterization of the proline/ecdione uptake system, Prop, and the ecdione/proline/glycine betaine carrier, EcpP," <i>J. Bacteriol.</i> , 180(22):6005-6012 (1998)
Y13221	glnA	Glutamine synthetase 1	Jakoby, M. et al. "Isolation of Corynebacterium glutamicum glnA gene encoding glutamin synthetase 1," <i>FEMS Microbial Lett.</i> , 154(1):81-88 (1997)
Y16642	lpd	Dihydrofolamide dehydrogenase	Moreau, S. et al. "Analysis of the integration functions of & phi;304L. An integrase module among corynephages," <i>Virology</i> , 253(1):150-159 (1999)
Y18039		Attachment site Corynephage 304L	Oguiza, J. A. et al. "A gene encoding arginyl-tRNA synthetase is located in the upstream region of the lysA gene in Brevibacterium lactofermentum. Regulation of argS-lysA cluster expression by arginine," <i>J. Bacteriol.</i> , 175(22):7356-7362 (1993)
221501	argS; lysA	Arginyl-tRNA synthetase; diaminopimelate decarboxylase (partial)	Pisabano, A. et al. "A cluster of three genes (dapA, orf2, and dapB) of Brevibacterium lactofermentum encodes dihydriodipicolinate reductase, and a third polypeptide of unknown function," <i>J. Bacteriol.</i> , 175(9):2743-2749 (1993)
221502	dapA; dapB	Dihydriodipicolinate synthase; dihydriodipicolinate reductase	Malumbres, M. et al. "Analysis and expression of the thrC gene of the encoded threonine synthase," <i>Appl. Environ. Microbiol.</i> , 60(7):2209-2219 (1994)
229563	thrC	Threonine synthase	Oguiza, J. A. et al. "Multiple sigma factor genes in Brevibacterium lactofermentum. Characterization of sigA and sigB," <i>J. Bacteriol.</i> , 178(2):550-553 (1996)
246753	16S rDNA	Gene for 16S ribosomal rRNA	Oguiza, J. A. et al. "The galE gene encoding the UDP-galactose 4-epimerase of Brevibacterium lactofermentum is coupled transcriptionally to the dmdR gene," <i>Gene</i> , 177:103-107 (1996)
249822	sigA	SigA sigma factor	Oguiza, J. A. et al. "Multiple sigma factor genes in Brevibacterium lactofermentum: Characterization of sigA and sigB," <i>J. Bacteriol.</i> , 178(2):550-553 (1996)
249823	galE; dmdR	Catalytic activity UDP-galactose 4-epimerase; diphtheria toxin regulatory protein	Caneja, A. et al. "Cloning and characterization of an IS-like element present in the genome of Brevibacterium lactofermentum ATCC 13869," <i>Gene</i> , 170(1):91-94 (1996)
249824	orf1; sigB	?; SigB sigma factor	
266534		Transposase	

A sequence for this gene was published in the indicated reference. However, the sequence obtained by the inventors of the present application is significantly longer than the published version. It is believed that the published version relied on an incorrect start codon, and thus represents only a fragment of the actual coding region.

TABLE 3: *Corynebacterium* and *Brevibacterium* Strains Which May be Used in the Practice of the Invention

GENUS	SPECIES	TYPE	DATE	NAME	REG. NO.	ISS. DATE
<i>Brevibacterium</i>	<i>ammoniagenes</i>		21034			
<i>Brevibacterium</i>	<i>ammoniagenes</i>		19350			
<i>Brevibacterium</i>	<i>ammoniagenes</i>		19351			
<i>Brevibacterium</i>	<i>ammoniagenes</i>		19352			
<i>Brevibacterium</i>	<i>ammoniagenes</i>		19353			
<i>Brevibacterium</i>	<i>ammoniagenes</i>		19354			
<i>Brevibacterium</i>	<i>ammoniagenes</i>		19355			
<i>Brevibacterium</i>	<i>ammoniagenes</i>		19356			
<i>Brevibacterium</i>	<i>ammoniagenes</i>		21055			
<i>Brevibacterium</i>	<i>ammoniagenes</i>		21077			
<i>Brevibacterium</i>	<i>ammoniagenes</i>		21553			
<i>Brevibacterium</i>	<i>ammoniagenes</i>		21580			
<i>Brevibacterium</i>	<i>ammoniagenes</i>		39101			
<i>Brevibacterium</i>	<i>butanicum</i>		21196			
<i>Brevibacterium</i>	<i>divaricatum</i>		21792	"928		
<i>Brevibacterium</i>	<i>flavum</i>		21474			
<i>Brevibacterium</i>	<i>flavum</i>		21129			
<i>Brevibacterium</i>	<i>flavum</i>		21518			
<i>Brevibacterium</i>	<i>flavum</i>			B11474		
<i>Brevibacterium</i>	<i>flavum</i>			B11472		
<i>Brevibacterium</i>	<i>flavum</i>		21127			
<i>Brevibacterium</i>	<i>flavum</i>		21128			
<i>Brevibacterium</i>	<i>flavum</i>		21427			
<i>Brevibacterium</i>	<i>flavum</i>		21475			
<i>Brevibacterium</i>	<i>flavum</i>		21517			
<i>Brevibacterium</i>	<i>flavum</i>		21526			
<i>Brevibacterium</i>	<i>flavum</i>		21529			
<i>Brevibacterium</i>	<i>flavum</i>			B11477		

<i>Brevibacterium</i>	<i>flavum</i>			B11478
<i>Brevibacterium</i>	<i>flavum</i>	21127		
<i>Brevibacterium</i>	<i>flavum</i>			B11474
<i>Brevibacterium</i>	<i>lcali</i>	15527		
<i>Brevibacterium</i>	<i>keioglutamicum</i>	21004		
<i>Brevibacterium</i>	<i>keioglutamicum</i>	21089		
<i>Brevibacterium</i>	<i>ketosorediculum</i>	21914		
<i>Brevibacterium</i>	<i>lactofermentum</i>		70	
<i>Brevibacterium</i>	<i>lactofermentum</i>			74
<i>Brevibacterium</i>	<i>lactofermentum</i>			77
<i>Brevibacterium</i>	<i>lactofermentum</i>	21798		
<i>Brevibacterium</i>	<i>lactofermentum</i>	21799		
<i>Brevibacterium</i>	<i>lactofermentum</i>	21800		
<i>Brevibacterium</i>	<i>lactofermentum</i>	21801		
<i>Brevibacterium</i>	<i>lactofermentum</i>		B11470	
<i>Brevibacterium</i>	<i>lactofermentum</i>			B11471
<i>Brevibacterium</i>	<i>lactofermentum</i>	21086		
<i>Brevibacterium</i>	<i>lactofermentum</i>	21420		
<i>Brevibacterium</i>	<i>lactofermentum</i>	21086		
<i>Brevibacterium</i>	<i>lactofermentum</i>	31269		
<i>Brevibacterium</i>	<i>linens</i>	9174		
<i>Brevibacterium</i>	<i>linens</i>	19391		
<i>Brevibacterium</i>	<i>linens</i>	8377		
<i>Brevibacterium</i>	<i>paraflinohyticum</i>		11160	
<i>Brevibacterium</i>	<i>spec.</i>			717.73
<i>Brevibacterium</i>	<i>spec.</i>			717.73
<i>Brevibacterium</i>	<i>spec.</i>	14604		
<i>Brevibacterium</i>	<i>spec.</i>	21860		
<i>Brevibacterium</i>	<i>spec.</i>	21864		
<i>Brevibacterium</i>	<i>spec.</i>	21865		
<i>Brevibacterium</i>	<i>spec.</i>	21866		
<i>Brevibacterium</i>	<i>spec.</i>	19240		

<i>Corynebacterium</i>	<i>acetocidophilum</i>	21476
<i>Corynebacterium</i>	<i>acetocidophilum</i>	13870
<i>Corynebacterium</i>	<i>acetoglutamicum</i>	BI1473
<i>Corynebacterium</i>	<i>acetoglutamicum</i>	BI1475
<i>Corynebacterium</i>	<i>acetoglutamicum</i>	15806
<i>Corynebacterium</i>	<i>acetoglutamicum</i>	21491
<i>Corynebacterium</i>	<i>acetoglutamicum</i>	31270
<i>Corynebacterium</i>	<i>acetophilum</i>	RJ671
<i>Corynebacterium</i>	<i>ammoniagenes</i>	2399
<i>Corynebacterium</i>	<i>ammoniagenes</i>	6872
<i>Corynebacterium</i>	<i>ammoniagenes</i>	15511
<i>Corynebacterium</i>	<i>sujiokense</i>	21496
<i>Corynebacterium</i>	<i>elutamicum</i>	14067
<i>Corynebacterium</i>	<i>elutamicum</i>	39137
<i>Corynebacterium</i>	<i>elutamicum</i>	21254
<i>Corynebacterium</i>	<i>elutamicum</i>	21255
<i>Corynebacterium</i>	<i>elutamicum</i>	31830
<i>Corynebacterium</i>	<i>elutamicum</i>	13032
<i>Corynebacterium</i>	<i>elutamicum</i>	14305
<i>Corynebacterium</i>	<i>elutamicum</i>	15455
<i>Corynebacterium</i>	<i>elutamicum</i>	13058
<i>Corynebacterium</i>	<i>elutamicum</i>	13059
<i>Corynebacterium</i>	<i>elutamicum</i>	13060
<i>Corynebacterium</i>	<i>elutamicum</i>	21492
<i>Corynebacterium</i>	<i>elutamicum</i>	21513
<i>Corynebacterium</i>	<i>elutamicum</i>	21526
<i>Corynebacterium</i>	<i>elutamicum</i>	21543
<i>Corynebacterium</i>	<i>elutamicum</i>	13287
<i>Corynebacterium</i>	<i>elutamicum</i>	21851
<i>Corynebacterium</i>	<i>elutamicum</i>	21253
<i>Corynebacterium</i>	<i>elutamicum</i>	21514
<i>Corynebacterium</i>	<i>elutamicum</i>	21516
<i>Corynebacterium</i>	<i>elutamicum</i>	21299

<i>Corynebacterium</i>	<i>glutamicum</i>	21100
<i>Corynebacterium</i>	<i>glutamicum</i>	39884
<i>Corynebacterium</i>	<i>glutamicum</i>	21488
<i>Corynebacterium</i>	<i>glutamicum</i>	21649
<i>Corynebacterium</i>	<i>glutamicum</i>	21650
<i>Corynebacterium</i>	<i>glutamicum</i>	19223
<i>Corynebacterium</i>	<i>glutamicum</i>	13869
<i>Corynebacterium</i>	<i>glutamicum</i>	21157
<i>Corynebacterium</i>	<i>glutamicum</i>	21158
<i>Corynebacterium</i>	<i>glutamicum</i>	21159
<i>Corynebacterium</i>	<i>glutamicum</i>	21355
<i>Corynebacterium</i>	<i>glutamicum</i>	31808
<i>Corynebacterium</i>	<i>glutamicum</i>	21674
<i>Corynebacterium</i>	<i>glutamicum</i>	21562
<i>Corynebacterium</i>	<i>glutamicum</i>	21563
<i>Corynebacterium</i>	<i>glutamicum</i>	21564
<i>Corynebacterium</i>	<i>glutamicum</i>	21565
<i>Corynebacterium</i>	<i>glutamicum</i>	21566
<i>Corynebacterium</i>	<i>glutamicum</i>	21567
<i>Corynebacterium</i>	<i>glutamicum</i>	21568
<i>Corynebacterium</i>	<i>glutamicum</i>	21569
<i>Corynebacterium</i>	<i>glutamicum</i>	21570
<i>Corynebacterium</i>	<i>glutamicum</i>	21571
<i>Corynebacterium</i>	<i>glutamicum</i>	21572
<i>Corynebacterium</i>	<i>glutamicum</i>	21573
<i>Corynebacterium</i>	<i>glutamicum</i>	21579
<i>Corynebacterium</i>	<i>glutamicum</i>	19049
<i>Corynebacterium</i>	<i>glutamicum</i>	19050
<i>Corynebacterium</i>	<i>glutamicum</i>	19051
<i>Corynebacterium</i>	<i>glutamicum</i>	19052
<i>Corynebacterium</i>	<i>glutamicum</i>	19053
<i>Corynebacterium</i>	<i>glutamicum</i>	19054

<i>Corynebacterium</i>	<i>glutamicum</i>	19055	
<i>Corynebacterium</i>	<i>glutamicum</i>	19056	
<i>Corynebacterium</i>	<i>glutamicum</i>	19057	
<i>Corynebacterium</i>	<i>glutamicum</i>	19058	
<i>Corynebacterium</i>	<i>glutamicum</i>	19059	
<i>Corynebacterium</i>	<i>glutamicum</i>	19060	
<i>Corynebacterium</i>	<i>glutamicum</i>	19185	
<i>Corynebacterium</i>	<i>glutamicum</i>	13286	
<i>Corynebacterium</i>	<i>glutamicum</i>	21515	
<i>Corynebacterium</i>	<i>glutamicum</i>	21527	
<i>Corynebacterium</i>	<i>glutamicum</i>	21544	
<i>Corynebacterium</i>	<i>glutamicum</i>	21492	
<i>Corynebacterium</i>	<i>glutamicum</i>	B8183	
<i>Corynebacterium</i>	<i>glutamicum</i>	B8182	
<i>Corynebacterium</i>	<i>glutamicum</i>	B12416	
<i>Corynebacterium</i>	<i>glutamicum</i>	B12417	
<i>Corynebacterium</i>	<i>glutamicum</i>	B12418	
<i>Corynebacterium</i>	<i>glutamicum</i>	B11476	
<i>Corynebacterium</i>	<i>glutamicum</i>		
<i>Corynebacterium</i>	<i>glutamicum</i>	21608	
<i>Corynebacterium</i>	<i>lilium</i>	1973	
<i>Corynebacterium</i>	<i>mirabilis</i>	21419	
<i>Corynebacterium</i>	<i>mirabilis</i>		
<i>Corynebacterium</i>	<i>spec.</i>	P4445	
<i>Corynebacterium</i>	<i>spec.</i>	P4446	
<i>Corynebacterium</i>	<i>spec.</i>	31088	
<i>Corynebacterium</i>	<i>spec.</i>	31089	
<i>Corynebacterium</i>	<i>spec.</i>	31090	
<i>Corynebacterium</i>	<i>spec.</i>	31090	
<i>Corynebacterium</i>	<i>spec.</i>	31090	20145
<i>Corynebacterium</i>	<i>spec.</i>	15954	
<i>Corynebacterium</i>	<i>spec.</i>	21857	
<i>Corynebacterium</i>	<i>spec.</i>	21862	
<i>Corynebacterium</i>	<i>spec.</i>	21863	

ATCC: American Type Culture Collection, Rockville, MD, USA

FERM: Fermentation Research Institute, Chiba, Japan

NRRL: ARS Culture Collection, Northern Regional Research Laboratory, Peoria, IL, USA

CECT: Colección Española de Cultivos Típicos, Valencia, Spain

NCIMB: National Collection of Industrial and Marine Bacteria Ltd., Aberdeen, UK

CBS: Centraalbureau voor Schimmelcultures, Baarn, NL

NCTC: National Collection of Type Cultures, London, UK

DSMZ: Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig, Germany

For reference see Sugawara, H. et al. (1993) World directory of collections of cultures of microorganisms: Bacteria, fungi and yeasts (4th edn), World Federation for culture collections world data center on microorganisms, Saitama, Japan.

Appendix A & B

>>RXA01992-amino acid sequence
(1-450, translated) 150 residues

VYCPFCQHDH SKVIDSRVID AGSAIRRRE CSKCEGRFTT IEKAVLLVVK RNGVTEPFSR EKVVTGVRRA
CQGRDVSSDA LKRLAQQVEE TVRSNGSSQV RANDIGLAIL DPLRELDEVA YLRFASVYKS FDSADDFEKE
IRLMRRGRD

>RXA01992-nucleotide sequence A: upstream

ATAGATCAATGCTGGCATAAACTAAAGGCACAAACAGCCGATTCACTTATTTTACCGAATGCTGAAGTCGAAGTA
AGTGAGGGGGAGAGGATCCGTTA

>RXA01992-nucleotide sequence B: coding region

GTGTACTGCCATTGCCAACATGATCATTCAAAAGTCATTGACTCCCGCGTCATTGACGCCGGAAGCGCCATTG
CAGGCCTCGAGTCAGCAAATGCGAAGGCCGTTCACCAACATCGAAAAAGCTGTTCTCCTCGTTAAAGAA
ACGGCGTCACTGAACCGTTAGTCGAGAAAAAGTAGTCACCGGTGTCGTCGTGCATGCCAAGGCCGCGACGTATCA
GATGACGCCTGAAACGCCACTGCTCACCAAGTCCAAGTCCAAGGAAAGCTCTCAAGTACCGCCTAA
CGATATTGGTTAGCCATTCTCGATCCACTGAGAGAACTCGACGAGGTAGCGTACCTACGCTTGCCCTGTGTATA
AGTCTTTGACAGTGCTGACGACTTGAAAAAGAAATCCGCCTCATGCGCAGACGCCGGAAGGGAC

>RXA01992-nucleotide sequence C: downstream

TAGCAGACACAAAGTCTTAACTA

Appendix A & B

>>RXA01942-amino acid sequence
(1-600, translated) 200 residues

MLRIGLTGGI GSGKSTVADL LSSEGFLIVD ADQVARDIVE PGQPALAEIA EAFGQDILKP DGTLDRAGLA
AKAFVSEEQT ALLNAITHPR IAEEESARRFN EAEDQGAKVA VYDMPPLLVEK GLDRKMDLVV VVDVDVEERV
RRLVEKRGLT EDDVRRRIAS QVPDDVRLKA ADIVVDNNGT LEDLHAEASK LIAEILSRVN

>RXA01942-nucleotide sequence A: upstream

GCCGCGAAATTCTGGTAAATTGAAGGTATTCTGCAGATCAGGCAAATTCTTCCACGACTGTGATCAAGGTCAACGG
CAAGAACGAGTAACCTGGGATCC

>RXA01942-nucleotide sequence B: coding region

ATGTTGCGCATTGGACTAACAGGAGGGATCGGCAGCGTAAATCTACCGTTGCCGATCTTTGTCACTGAAGGATT
TCTCATCGTCACGCCAACAGTTGCCGCGATATCGTCAACCCGGACAACCGGCATTAGCAGAGCTAGCTGAAG
CTTTTGGCCAAGACATCTAAAACCCGACGGCACTCTAGACCGCGCGGGATTAGCAGCCAAGCATTGTCAAGCGAA
GAACAAACAGCGCTGCTCAATGCCATTACCCACCCCTCGTATCGCGAAGAGTCAGCTCGTCGATTCAACGAAGCCGA
AGATCAAGGCGCAAAGTTGGTTTATGACATGCCTTGCTTAGAAAAAGGCCCTGACCGCAAGATGGACCTTG
TCGTCGTAGTTGATGTTGACGTAGAGGAACCGCTCCGCAGACTTGTGGAAAAACGTGGCCCTCACAGAGGACGACGTG
CGGCGTCAATCGCTTCTCAAGTGCCGACGACGTCAAGCTAAAGCCGCTGACATCGTTGACGAAATAACGGCAC
GCTAGAGGACCTTCATGCTGAAGCAAGCTGATTGCTGAGATTCTTAGTCGGTGAAT

>RXA01942-nucleotide sequence C: downstream

TAGCACTAAAACATCGTCAAAGT

Appendix A & B

>>RXA01931-amino acid sequence
(1-654, translated) 218 residues

MSTSFESIPG VTISARKALS TAGFKDLESL AGTNYEEVAG LSGIGARTLE RLQAALVEKR MSFGGVPEA
EQRTATWTTL DSAAPEATET SESPEYFIQN LDIPRRITHG RLLLEIFNRA TGQKPYVAGS SIVGYGRVHY
RYATGREGIT IRVGFSPRKA KISLYGLTSA PASRELLKKL GKHSVGVSCL YINKPEDVDL EVLEEMIRIS
WEAEPGEC

>RXA01931-nucleotide sequence A: upstream

TCGATAGGCATGGCGCCGATTCTAGGTCGAAGTCTAAAGTCCCACCGCAAAAGTAACCAAGGCGTACATCTA
CCCGGCACCTTAAGATCAGTGGC

>RXA01931-nucleotide sequence B: coding region

ATGAGCACTTCTTGAGTCCATTCCCGGAGTTACTATCTCCGCTCGAAAAGCCTTGAGCACTGCGGGTTAAAGA
TCTGGAATCGCTTGCAGGTACCAATTACGAAGAAGTGGCTGGTCTGTCAGGAATTGGCGCACGTACTTGGAGCGAT
TGCAAGCAGCGCTGGTGGAGAAGAGAATGAGCTTGGCGGTAAAGTGCCAGAGGCAGAACAGCGCACCACGTGG
ACAACGCTAGATTCTGCAGCACCAAGAGGCAACAGAAACCTCAGAATCCCCAGAGTACTTCATTCAAAACCTAGACAT
TCCTCGCAGAATTACCCACGGCCGGTTGTTGGAAATTTCACACAGAGCAGGCCAGAAACCTTATGTAGCAG
GGTCTTCGATTGTTGGTATGGCGAGTTCAATTATCGCTATGCCACGGTAGGGAAGGGATCACTATTCTGTGGGT
TTCAGTCCCGCAAAGGAAAGATTTCACTTTATGGACTAACAGTGCCTGCCTCGAGAACTTCTGAAGAAATT
GGGTAAGCATTCTGGTGTGCCTGCCTATATCAACAAACCAAGAGGATGTTGATCTGAGGTGTTGAGGAAA
TGATCCGTATTCTGGGAAGCGGAACCTGGCGAATGT

>RXA01931-nucleotide sequence C: downstream

TAAATCTCAACCTTAACAAAGT

Appendix A & B

>>RXA01916-amino acid sequence
(1-1011, translated) 337 residues

MKLGLYNAIF HDRTLPEALA AIKAAGLTGI ELNTGGFLPA THIPTIDDIL VSDDARDEFI GIFE GTGVDI
YGLNCNGNPL HPNKAIGDKH AEDIRRSIRL AERLGQNRVV TMSGLPGGE P GAKYTNWVNV AWNSAALDVL
DYQWDIAAEF WRETDRAFAAD HGVKVALELH PQNIVFNSAD VHKLIDL TGA THVGVELDAS HLFWQQMDPI
AVIDHLGELI FHAAAKDVRV NKEWAQLNGV LDNSFRRLD P SENRTNLGGD EWANEWPKNS AWDFVALGRG
HDVAYWTEFL RALHRVDPNM LVNIEHEDVS LGREEGVNEA AKVLIEANKA LEESLVS

>RXA01916-nucleotide sequence A: upstream
GCTGTATCACAGTCAGCTGCAGCAGGTGGCGAAACCGTTGCGGTCCCAGCGGCTGCTCTGATCCCTGCAAACAACTA
GAAACTATTAGAAAGCATCACC

>RXA01916-nucleotide sequence B: coding region
ATGAAACTCGGTCTCTACAACCGCAGTCCTCCACCGACCGCACCCCTGCCAGAAGCGCTCGCAGCCATCAAAGCTGCAGG
TCTCACCGGAATTGAACTCAACACCGCGGATTTCGCCTGCAACCCACATCCCGACCATCGATGACATCCTGGTCA
GGCATGATGCCCGCGATGAATTCTCGGGATTTCGAAGGCACCGGGCTCGACATCTACGGCCTTAAC TGCAACGGC
AACCCGCTTCACCCCAACAAGCGATCGGGGACAAGCATGCCGAAGACATTGACGTTCCATCCGCCCTCCGAGAGCG
CCTCGGCCAAAACCGTGTGGTCACCATGTCTGGCTTCTGGTGGCGAACCGGGCGAACGTACACCAACTGGGTTG
TCAACCGCGTGGAACTCCGAGCCTGGATGTCCTTGATTACCAATGGGATATCGCAGCTGAATTCTGGCGCGAGACC
GACCGCTTCGCCCGAGATCACGGCGTAAAGTGGCTCTTGAGCTGCACCCACAGAACATCGTGTCAACTCCGCTGA
CGTGCATAAGCTCATCGATCTCACCGCGCCACCCACGTGGCGTGAACCTGGATGCATCACACCTGTTCTGGCAGC
AGATGGACCCAATCGCTGTGATTGATCACCTCGCGAGCTCATCTCCACGCCGCCAAAGACGTGCGAGTTAAT
AAGGAATGGGCTCAGCTCAACGGTGTGGACAAACAGCTCCGACGCCTGACCCATCCGAAAACCGCACCAACTT
GGCGGGCGACGAGTGGCGAATGAATGGCAAAGAAACTCTGCTTGGGATTCTGTTGCTCTGGGCCGGTCATGACG
TTGCTTACTGGACCGAATTCTCCCGCGACTTCACCGCGTCACTTCAACATGCTGGTCAACATCGAACACGAGGAT
GTTTCACTCGGTGCGAAGAAGCGTCAACGAAGCCGCTAAGGTGCTGATCGAGGCCAACAGGCACACTGAAGAGTC
CCTGGTTTCT

>RXA01916-nucleotide sequence C: downstream
TAAAAAAACTACGCCCTGCCCGC

Appendix A & B

>>RXA01896-amino acid sequence
(1-579, translated) 193 residues

ARAALHAWSP TVGVDTDVER WIELDKWGFA RFERGETTHL GQRDRIRAY LNRELDDATC DDIYSGYLKA
YEQNWTAYPD AKGVLDRAVA TGAPVGILTN GAAPMQQDKL DRTGLGLPEL VMLAATLDS AKPRPEMYAR
ALTHLGARTA TIIGDDWTND VAAPRELGWN ALYLDRSGTD PRADIHSLDE LFH

>RXA01896-nucleotide sequence B: coding region

GCCCCGCCGCCCTGCACGCTGGTCGGCGACCGTGGCGTCGACACGGATGTTGAGCGCTGGATTGAGCTGGATAA
GTGGGGTTTGGCCGTTTGAGCGCGGAAACCAACGCATTAGGTCAAGCGCGACCGCATTAGGCATCAGGGCGTACCTCA
ACAGAGAGCTTGACGACGCCACCTGCGATGATATTACTCCGGTACCTTAAAGCATATGAGCAAAACTGGACTGCC
TACCCCCGATGCCAAGGGCGTCTCGATCGCGCGGTAGCCACCGGTCCCCCTGTGGGAATCCTGACCAATGGCGCAGC
CCCCATGCAGCAAGACAAGCTTGATCGCACCGGCCTTGGCCTGCCAGAACTCGTCATGTTGGCGCGTCCACTCTGG
ATTCTGCGAAGCCTCGCCCCGAAATGTATGCCGAGCGCTACCCATTGGGTGCCGAACCGCAACAATTATCGGC
GATGATTGGACCAACGATGTCGAGCTCCCCGCGAACTTGGCTGGAATGCTCTATTAGATCGTCCGGAAACCGA
TCCACGCGCCGATATCCACTCCCTGGATGAACCTTTCACT

>RXA01896-nucleotide sequence C: downstream

TAGGCTGGCCTTATTGTTCCG

Claims

1. An isolated nucleic acid molecule from *Corynebacterium glutamicum* encoding an MCP protein, or a portion thereof.
5
2. The isolated nucleic acid molecule of claim 1, wherein said nucleic acid molecule encodes an MCP protein involved in fine chemical production.
3. An isolated *Corynebacterium glutamicum* nucleic acid molecule selected from the group consisting of those sequences set forth in Appendix A, or a portion thereof.
10
4. An isolated nucleic acid molecule which encodes a polypeptide sequence selected from the group consisting of those sequences set forth in Appendix B.
5. An isolated nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide selected from the group of amino acid sequences consisting of those sequences set forth in Appendix B.
15
6. An isolated nucleic acid molecule comprising a nucleotide sequence which is at least 50% homologous to a nucleotide sequence selected from the group consisting of those sequences set forth in Appendix A, or a portion thereof.
20
7. An isolated nucleic acid molecule comprising a fragment of at least 15 nucleotides of a nucleic acid comprising a nucleotide sequence selected from the group consisting of those sequences set forth in Appendix A.
25
8. An isolated nucleic acid molecule which hybridizes to the nucleic acid molecule of any one of claims 1-7 under stringent conditions.
9. An isolated nucleic acid molecule comprising the nucleic acid molecule of any one of claims 1-8 or a portion thereof and a nucleotide sequence encoding a heterologous polypeptide.
30
10. A vector comprising the nucleic acid molecule of any one of claims 1-9.
35
11. The vector of claim 10, which is an expression vector.
12. A host cell transfected with the expression vector of claim 11.
13. The host cell of claim 12, wherein said cell is a microorganism.
40
14. The host cell of claim 13, wherein said cell belongs to the genus *Corynebacterium* or *Brevibacterium*.
15. The host cell of claim 12, wherein the expression of said nucleic acid molecule results in the modulation in production of a fine chemical from said cell.
45

16. The host cell of claim 15, wherein said fine chemical is selected from the group consisting of: organic acids, proteinogenic and nonproteinogenic amino acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated and unsaturated fatty acids, diols, carbohydrates, aromatic compounds, vitamins, cofactors, and enzymes.

5 17. A method of producing a polypeptide comprising culturing the host cell of claim 12 in an appropriate culture medium to, thereby, produce the polypeptide.

10 18. An isolated MCP polypeptide from *Corynebacterium glutamicum*, or a portion thereof.

15 19. The polypeptide of claim 18, wherein said polypeptide is involved in fine chemical production.

15 20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of those sequences set forth in Appendix B.

20 21. An isolated polypeptide comprising a naturally occurring allelic variant of a polypeptide comprising an amino acid sequence selected from the group consisting of those sequences set forth in Appendix B, or a portion thereof.

25 22. The isolated polypeptide of any of claims 18-21, further comprising heterologous amino acid sequences.

25 23. An isolated polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 50% homologous to a nucleic acid selected from the group consisting of those sequences set forth in Appendix A.

30 24. An isolated polypeptide comprising an amino acid sequence which is at least 50% homologous to an amino acid sequence selected from the group consisting of those sequences set forth in Appendix B.

35 25. A method for producing a fine chemical, comprising culturing a cell containing a vector of claim 12 such that the fine chemical is produced.

26. The method of claim 25, wherein said method further comprises the step of recovering the fine chemical from said culture.

40 27. The method of claim 25, wherein said method further comprises the step of transfecting said cell with the vector of claim 11 to result in a cell containing said vector.

45 28. The method of claim 25, wherein said cell belongs to the genus *Corynebacterium* or *Brevibacterium*.

29. The method of claim 25, wherein said cell is selected from the group consisting of: *Corynebacterium glutamicum*, *Corynebacterium herculis*, *Corynebacterium lilium*, *Corynebacterium acetoacidophilum*, *Corynebacterium acetoglutamicum*.

Corynebacterium acetophilum. *Corynebacterium ammoniogenes*. *Corynebacterium fujiiokense*. *Corynebacterium nitrilophilus*. *Brevibacterium ammoniagenes*.
5 *Brevibacterium butanicum*. *Brevibacterium davaricatum*. *Brevibacterium flavum*.
Brevibacterium healii. *Brevibacterium ketoglutamicum*. *Brevibacterium ketosoreductum*. *Brevibacterium lactofermentum*. *Brevibacterium linens*.
Brevibacterium paraffinolyticum. and those strains set forth in Table 3.

30. The method of claim 25, wherein expression of the nucleic acid molecule from said vector results in modulation of production of said fine chemical.

10 31. The method of claim 25, wherein said fine chemical is selected from the group consisting of: organic acids, proteinogenic and nonproteinogenic amino acids, purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated and unsaturated fatty acids, diols, carbohydrates, aromatic compounds, vitamins, cofactors, and enzymes.

15 32. The method of claim 25, wherein said fine chemical is an amino acid.

20 33. The method of claim 32, wherein said amino acid is drawn from the group consisting of: lysine, glutamate, glutamine, alanine, aspartate, glycine, serine, threonine, methionine, cysteine, valine, leucine, isoleucine, arginine, proline, histidine, tyrosine, phenylalanine, and tryptophan.

25 34. A method for producing a fine chemical, comprising culturing a cell whose genomic DNA has been altered by the inclusion of a nucleic acid molecule of any one of claims 1-9.

CORYNEBACTERIUM GLUTAMICUM GENES ENCODING NOVEL PROTEINS

Abstract of the Disclosure

5

Isolated nucleic acid molecules, designated MCP nucleic acid molecules, which encode novel MCP proteins from *Corynebacterium glutamicum* are described. The invention also provides antisense nucleic acid molecules, recombinant expression vectors containing MCP nucleic acid molecules, and host cells into which the expression vectors have been introduced. The invention still further provides isolated MCP proteins, mutated MCP proteins, fusion proteins, antigenic peptides and methods for the improvement of production of a desired compound from *C. glutamicum* based on genetic engineering of MCP genes in this organism.

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.